Characterisation of stem rust resistance genes effective against race Ug99 on wheat chromosome 7L

Vincent Pujol

A thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

of The Australian National University

Research School of Biology

Division of Plant Sciences

THE AUSTRALIAN NATIONAL UNIVERSITY

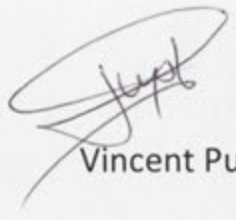
Canberra ACT 0200

Australia

June 2014

Declaration by the author

The work presented in this thesis is original, and to the best of my knowledge, it does not contain any material previously published or written by another person, except where due references are made in the text. Specific contributions by others are duly mentioned in the text. This thesis has not been submitted for any other degree.



Vincent Pujol

Acknowledgments

It is with great pleasure that I offer my sincerest gratitude to my supervisors, Dr Evans Lagudah and Dr Linda Tabe, for giving me the opportunity to undertake my PhD studies and for their valuable guidance throughout the program. I am thankful to Assoc. Prof David Jones, of the Australian National University, for his academic supervision. I thank our collaborators, Dr Li Huang and Dr. Sridhar Bhavani, for their contribution to the project, as well as the BREAD program of the National Science Foundation for financial support.

I am truly indebted to Prof Robert McIntosh of the Plant Breeding Institute at the University of Sydney, who taught me the key notions of rust phenotyping, as well as to Drs. Peng Zhang and Sami Hoxha for kindly coordinating the phenotyping experiments. I am grateful to Dr Matthew Rouse from the Department of Plant Pathology at the University of Minnesota, for the Ug99 phenotyping.

This thesis would not have been possible without the help of Dr Matthew Hayden of the Victorian Department of Primary Industries and his team for coordinating the SNP chip and genotyping by sequencing experiments. Likewise, I thank Dr. Jen Taylor of CSIRO and her team for their help in RNA-Seq and providing access to informatic resources.

I am extremely grateful to all the team members, Dr Sambasivam Periyannan, Mrs Sutha Chandramohan and Dr Wendelin Schnippenkoetter, to name but a few, who took the time to train me. A special thanks to my colleagues and friends Philip Watkins and Bosibori Bett, whose knowledge of the English language was of great worth. I further thank all colleagues at CSIRO Plant Industry for their precious advice, support and friendship.

Finally, but importantly, my most profound thanks go to my beloved companion Amrit, for her endless support and for always being by my side.

Abstract

Wheat is one of the most important food crops in the world, grown on more acreage than any other. *Puccinia graminis* f. sp. *tritici*, the causal agent of wheat and barley stem rust, is responsible for major production losses around the world. The development of resistant cultivars is an effective way to manage the disease, but outbreaks can occur when new pathogen races overcome the existing resistances. This happened with Ug99 (TTKSK), a highly virulent race now spreading in Africa and the Middle East and dangerously threatening worldwide food security. Thus, constant efforts to find new sources of resistance must remain a priority.

The hexaploid wheat cultivar Canthatch has long been known for carrying a suppressor of stem rust resistance on chromosome 7DL. This suppressor is now of particular interest, as the genes it inhibits were found to be effective against Ug99. Because the 7DL suppressor seems to be common in wheat cultivars, cloning it would be the first step toward unlocking masked resistance genes, as well as providing new information about the nature of action of a previously unstudied class of negative regulators of plant defence responses. This study first aimed to genetically map the 7DL suppressor.

A $F_{2:3}$ mapping population was developed from the cross between the susceptible line Columbus, thought to possess the 7DL suppressor, with the resistant line Columbus-NS, a near-isogenic line in which a mutated allele of the suppressor was inferred to have been transferred into the Columbus genetic background. The mapping population was phenotyped using two stem rust races, Ug99 (TTKSK) and one of Australian origin (#313), for which results were similar, indicating a common resistance against both races. Molecular markers were found using two recently developed techniques: i) genotyping of nearly 9,000 putative single nucleotide polymorphisms (SNP) using the wheat 9k SNP chip and ii) genotyping by sequencing using the restriction-site associated DNA sequencing (RAD-Seq) method in association with bulked segregant analysis. This led to the rapid identification of thousands of markers putatively linked to the stem rust resistance in Columbus-NS.

Surprisingly, genetic mapping revealed the resistance to be located on 7AL, not 7DL, indicating that it was due to a new resistance locus rather than a non-suppressor. This locus is most likely part of the resistance complex repressed by the 7DL suppressor. Although segregation ratios in the $F_{2:3}$ population indicated monogenic segregation (1:2:1), phenotypic observations and genomic studies indicated the possible involvement of multiple genes in the

resistance, notably on chromosome 6A. The 7AL locus seemed to be essential but not sufficient for full resistance.

The resistance response was also investigated at a microscopic level, which showed the involvement of cell death during the rust penetration through the stoma, at two days post-inoculation. Transcriptome analysis using RNA-Seq revealed 353 genes differentially expressed between Columbus and two independent Columbus-NS lines. Notably, several genes coding for cysteine-rich receptor-like kinases were found to be good candidates for involvement in the defence response.

Table of content

Chapter 1

Introduction

I. Wheat

I.1. Crop

Wheat (*Triticum* spp.) is a cereal and one of the most important food crops in the world. It represents the largest crop acreage worldwide, occupying 17% of the total crop area and is grown from Norway to Argentina (FAOSTAT 2013, http://faostat.fao.org/). It is the 4th most produced crop with over 650 millions tons harvested each year, behind sugar cane and two other cereals, maize and rice. In many regions, wheat is at the very basis of human nutrition and one of the main food supplies, especially in less developed countries where it is a substantial part of the diet. With a constant increase in world food demand due to an ever-growing population, expected to reach between 8.3 and 10.9 billion by 2050 (UN 2012, http://www.un.org/), wheat appears as a crucial factor for maintaining food security.

I.2. Domestication

Historically, wheat was a founding element in human civilizations. It is one of the earliest species to have been domesticated, first cultivated during the Neolithic period about 10,000 years ago in south-eastern of Turkey (Heun, 1997; Shewry, 2009). Over the centuries, wheat went through many modifications due to selection pressure, which changed it from a wild grass to a highly productive species. Two major changes included the loss of spikelet shattering at maturity, preventing seed loss at harvesting, and glume reduction, which facilitated threshing (Gill et al., 2007).

I.3. Genomic

At a genomic level, the wheat group has undergone several events of allopolyploidization, the hybridization of different species, and can now be found at several levels of ploidy (review: Matsuoka, 2011). Cultivated wheat comprises:

- Diploid species: Einkorn wheat (*T. monococcum*: AA, 2n=14). One of the earliest cultivated forms of wheat.
- Tetraploid species: Durum or pasta wheat (*T. durum*: AABB, 2n=28). Originated from the hybridization between two diploid species: *T.urartu* (AA, 2n=14), the donor of the A genome (Chapman et al., 1976; Dvorak, 1976; Dvorak et al., 1993) and

an undiscovered *Aegilops* species related to *Ae. speltoides* (SS, 2n=14), the donor of the B genome (review: Haider, 2013).

- Hexaploid species: Bread wheat (*T. aestivium*: AABBDD, 2n=42). Originated from the hybridization of the tetraploid *T. turgidum* (AABB, 2n=28) and the diploid *Ae. tauschii* (DD, 2n=14), the donor of the D genome (McFadden and Sears, 1946). Bread wheat accounts for 95% of the wheat grown, the remaining 5% being mainly durum.

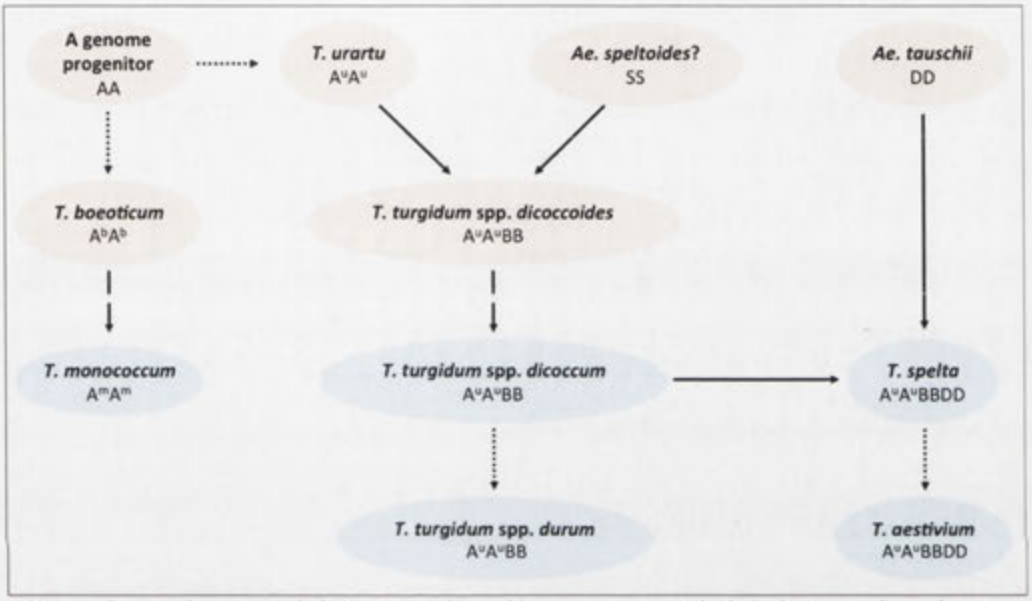


Figure 1: Wheat evolution. Fine dashed arrows indicate divergence events, wide dashed arrows indicate domestication and solid arrows indicate allopolyploidization events. Wild species are shown in orange and cultivated species in blue.

With its three homoeologous genomes (A, B and D), the bread wheat genome is one of the largest in higher plants (17 Gb) which is around 40 and 135 times larger than that of rice and *Arabidopsis*, respectively (Arabidopsis Genome Initiative, 2000; International Rice Genome Sequencing Project, 2005). Around 80% of the genome is composed of repetitive elements, mostly retrotransposons (Brenchley et al., 2012; Smith and Flavell, 1975). Because of its genome size and complexity, wheat genomics has always been challenging.

Although a reference genome is yet to be completed, significant progress has been made in the last few years. A low coverage whole-genome shotgun sequencing has already been published (Brenchley et al., 2012) and an international effort is being made to soon provide a reference of quality comparable to that of the rice genome (IWGSC,

http://www.wheatgenome.org/). Draft genomes of wheat A and D genome progenitors, *T. urartu* and *Ae. tauschii* respectively, are also available (Jia et al., 2013; Ling et al., 2013).

I.4. Diseases

Like any other plants, wheat is subject to various diseases. Some of the most destructive comprise rusts, powdery mildew, spot blotches and fusarium head blight. Epidemics can cause serious losses in crop yield, compromising food security and damaging the economy. In Australia, wheat diseases were estimated to cause an average loss of $913 million per year representing 19.5% of the value of production (Murray and Brennan, 2009). In countries with less developed agricultural practices, these losses may be higher. With 842 million people still being undernourished (FAO, 2013), disease control is critical for reducing world hunger and for sustaining food production.

Many solutions can help limiting the spread of diseases such as good agricultural practices, intercropping, crop rotation, plant quarantine and the use of resistant cultivars. Regardless of the means, understanding the interaction between plants and pathogens is crucial for efficient management of crop diseases.

II. Plant immunity

Plant diseases can be caused by a broad range of phytopathogens including viruses, bacteria, fungi, oomycetes and nematodes. Depending on their lifestyles, these pathogens can be divided into 3 main categories; biotrophs keep their host alive and hijack nutrients from living tissues whereas necrotrophs kill their host and feed on the release of nutrients from dead tissues. Hemibiotrophs initially are biotrophic but become necrotrophic at latter stages.

Preventing the entry of pathogens constitutes the first line of defence and is sufficient against the majority of microorganisms. This is mainly achieved with physical barriers such as the cell wall and layers of waxy cuticle (Reina-Pinto and Yephremov, 2009). Pathogens usually penetrate the plants by forcing their way through the leaf or root surface using specialized structures (e.g. appressorium), or directly through wounds or natural openings (e.g. stomata). In response, plants evolved two sophisticated defence mechanisms, namely, pattern-triggered immunity (PTI) and effector-triggered immunity (ETI). PTI is also referred to as microbial-associated molecular patterns-triggered immunity (MTI).

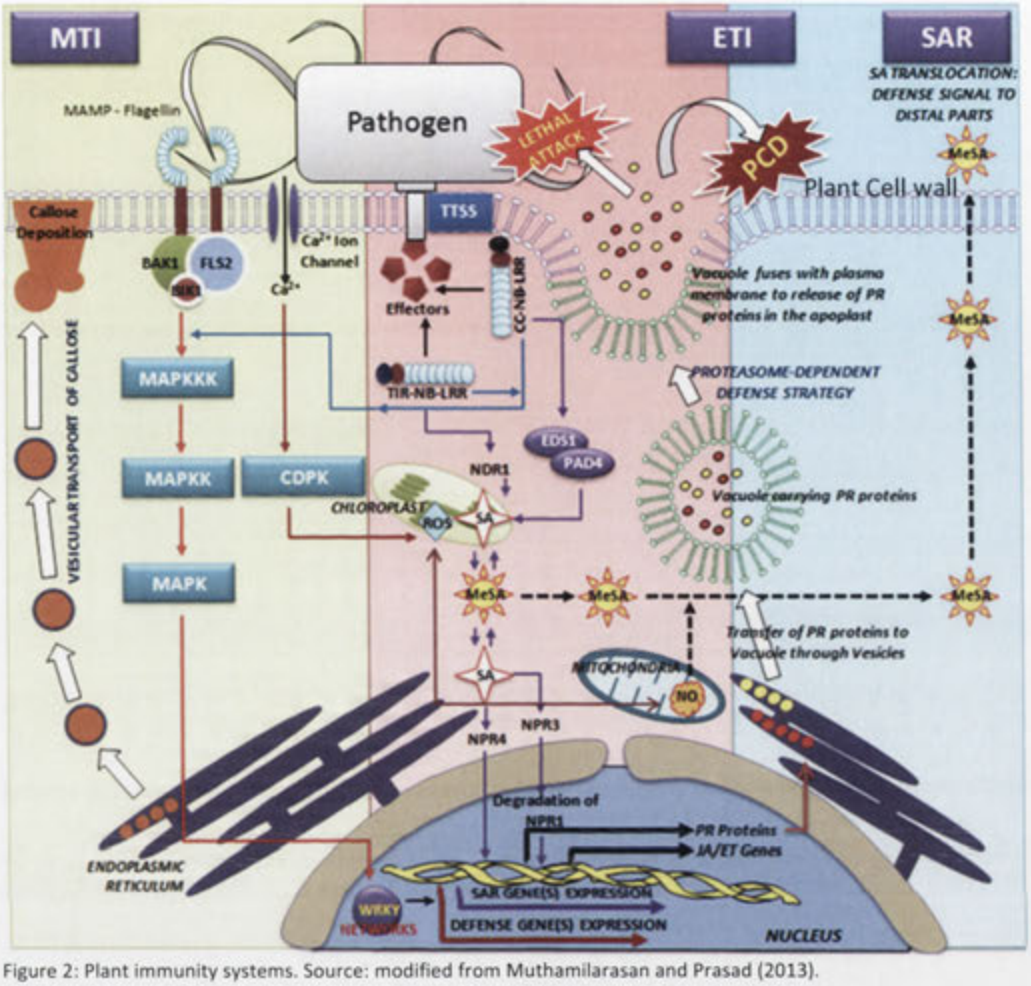


Figure 2: Plant immunity systems. Source: modified from Muthamilarasan and Prasad (2013).

II.1. Pattern-triggered immunity (PTI)

PTI is the first active defence mechanism and protects against a wide range of pathogens. Following induction of PTI, various defence responses are established, such as a burst of reactive oxygen species (ROS) (Chandra and Low, 1995), accumulation of callose at the sites of infection (Rodriguez-Galvez and Mendgen, 1995), stomatal closure (Melotto et al., 2006) and photosynthesis reduction (Göhre et al., 2012). These measures are efficient against all non-adapted pathogens.

MAMPs. PTI is induced through the recognition of microbial-associated molecular patterns (MAMPs) which are essential components of microorganisms, such as bacterial peptidoglycan (Gust et al., 2007) and flagellin (Felix et al., 1999), or fungal chitin (Felix et al., 1993) and ergosterol (Granado et al., 1995). Because MAMPs are necessary to their life, they are

strongly conserved throughout whole classes of microorganisms. Plants can also perceive host-derived degradation products, called damage-associated molecular patterns (DAMPs), which are released during infection.

PRRs. These molecular patterns are detected by a range of different pattern recognition receptors (PRRs) located in the plasma membrane (review: Monaghan and Zipfel, 2012). PRRs belong to two main protein families: receptor-like kinases (RLKs) or receptor-like protein (RLPs). Both receptors are composed of a transmembrane domain and an extracellular domain, generally a leucine-rich repeat (LRR) domain, for ligand binding. RLKs also possess a cytoplasmic Ser/Thr protein kinase domain, whereas RLPs only have a short cytoplasmic tail. Upon recognition, PRRs often heterodimerize with the LRR receptor kinase BRASSINOSTEROID INSENSITIVE 1-ASSOCIATED KINASE 1 (BAK1). The signal is then rapidly transduced, often through activation of mitogen-activated protein kinase (MAPK) cascades leading to the expression of defence-related genes.

Effectors. Adapted pathogens developed strategies to circumvent PTI and remain undetected. They release an array of virulence proteins called effectors, that facilitate establishment of the pathogen (Hogenhout et al., 2009). Effectors are highly diverse and play various roles, notably in the suppression of the host immune response. Some promote virulence by inhibiting plant perception of the pathogen; e.g. AvrPtoB from *P. syringae* inhibits flagellin recognition by interacting with the kinase domains of the PRR FLS2 as well as BAK1 (Göhre et al., 2008; Shan et al., 2008). Others alter host-cell structures; e.g. COR from *P. syringae* induces stomatal opening by blocking the MAMP signal transduction pathway (Melotto et al., 2006). Many effectors are directly delivered into the host cell by way of specialized structures: the type III secretion system (T3SS) in bacteria, the haustorium in fungi and oomycetes and the stylet in nematodes. Other effectors are secreted by the pathogen into the apoplast of invaded cells.

II.2. Effector-triggered immunity (ETI)

Just like effectors are the products of co-evolution between plants and pathogens, allowing the latter to modify specific host targets for their own benefits, plants developed tools to recognize these effectors. This recognition leads to the induction of ETI, the second active defence mechanism. ETI is typically characterised by a hypersensitive response (HR) at the site of attempted infection, stopping the pathogen from propagating further. HR is accompanied

by programmed cell death and the production of antimicrobial molecules such as the pathogenesis-related (PR) proteins (Mur et al., 2008; Spoel and Dong, 2012).

Unlike PTI, which provides moderate resistance and is induced by receptors that can perceive many microorganisms, the ETI response is strong, and receptors are specific to effectors from few adapted pathogens. The difference in the detection range is due to the element detected. While in PTI, PRRs recognize limited conserved MAMPs, ETI is based on the recognition of highly polymorphic effectors that greatly vary in pathogen populations (Dodds and Rathjen, 2010). As a consequence of the variety of effectors, their plant receptors, known as resistance (R) proteins, are numerous and highly specialized. The specific relationship between effectors and R proteins has led to the concept of "gene-for-gene" interaction (Flor, 1971), stating in its simplest form that one R protein specifically recognises one effector or more exactly, one avirulence (Avr) factor.

R proteins. As many effectors are released in the cell, most R proteins are intracellular and belong to the NB-LRR protein family, one of the largest gene families in plants (Marone et al., 2013). These proteins contain a C-terminal LRR domain, a central nucleotide binding (NB) domain and either a Coiled-Coil (CC) or a Toll/Interleukin-1 Receptor (TIR) domain in N-terminal. Although the overall structure of NB-LRR is conserved across plant species, the LRR domain is highly variable, reflecting their role in the recognition of specific effectors.

Recognition. NB-LRR proteins can recognize effectors in two different ways (Dodds and Rathjen, 2010). The first is direct and involves the physical interaction between effector and receptor. In the second type, the effector is indirectly recognized through an intermediate protein. Several models of indirect recognition have been proposed and differ in the role of the intermediate protein. In the guard model, the effector targets and modifies the intermediate protein, or guardee, which is kept under surveillance by the NB-LRR protein. The decoy model is similar to the guard model, but the intermediate protein mimics the protein that is usually targeted by the effector, and do not have any other function. Finally, in the bait-and-switch model, the intermediate protein gathers effector and receptor together to facilitate direct interaction.

Interestingly, there are many similarities between plant and animal immune systems, notably the structure of the receptors (PRRs and R proteins) involved in pathogen recognition. These systems, however, are not related but have probably evolved in a convergent manner

(Ausubel, 2005; Ronald and Beutler, 2010). Contrary to plants, animals possess memory immune cells which confer long-lasting and more effective resistance after primary antigen recognition. Nevertheless, plants also developed long lasting protection, known as systemic acquired resistance (SAR).

II.3. Systemic acquired resistance (SAR)

As seen above, induction of PTI and ETI by an avirulent pathogen leads to the onset of various defence responses that are localised at or around the site of infection. These mechanisms also trigger the biosynthesis of several hormones, notably salicylic acid (SA), jasmonic acid (JA) and ethylene (ET). These hormones are often seen as antagonist due to the type of pathogens against which they provide resistance (Derksen et al., 2013). Indeed, SA is usually implicated in resistance against biotrophs whereas JA and ET are involved in resistance against necrotrophs. They often lead to the production of different PR proteins; PR1, PR2 and PR5 are induced by SA, whereas PR3, PR4 and PR12 by JA. However, the biosynthesis of these hormones in defence is not exclusive as they often cooperate through crosstalk between their respective pathways.

Local infections also lead to accumulation of SA in unharmed distal tissues and production of PR proteins (Fu and Dong, 2013). These events are characteristics of the SAR mechanism, which confers systemic resistance against a broad-spectrum of pathogens. Because it can last for weeks to months, SAR-induced resistance is often considered as the plant immune system memory.

III. Wheat stem rust

Rusts are basidiomycete fungi causing some of the most destructive diseases in plants. The genus *Puccinia* is the largest one, and members are serious pathogens of all major cereal crop species except rice (Ayliffe et al., 2011). Recurrent epidemics of cereal rusts have been observed since the dawn of agriculture, causing catastrophic famines and economic crisis (Roelfs et al., 1992). In wheat, rusts are caused by three agents:

- *Puccinia graminis* f. sp. *tritici* (*Pgt*) for stem or black rust,
- *Puccinia triticina* (*Pt*) for leaf or brown rust,
- *Puccinia striiformis* f. sp. *tritici* (*Pst*) for stripe or yellow rust.

Stem rust was once the most feared disease of cereal crops, due to its rapid spread on a very large scale and its capability of ruining an entire healthy crop within a month. Because of its extreme potential of destruction, measures to manage stem rust had been widely taken, notably though the development of resistant cultivars and the eradication of common barberry, the alternate host. Although under control now, outbreaks can still occur. This is currently happening with Ug99, a highly virulent race that is dangerously spreading in Africa and the Middle East. Constant efforts to develop new strategies to fight wheat stem rust must remain a priority in order to maintain food security.

III.1. Life cycle

Stem rust fungus is a heteroecious obligate biotroph with five spore stages and two hosts. Primary hosts include bread and durum wheat, barley (*Hordeum vulgare*), rye and triticale (× *Triticosecale*). Common barberry (*Berberis vulgaris*) is considered the main alternate host.

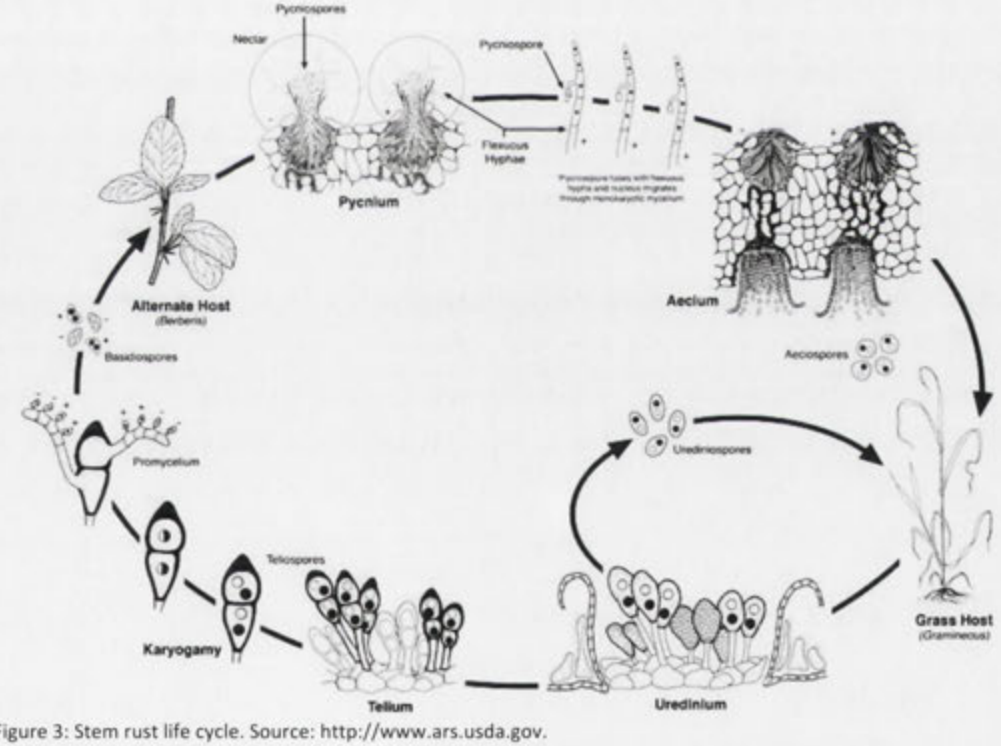


Figure 3: Stem rust life cycle. Source: http://www.ars.usda.gov.

In the full stem rust life cycle, Aeciospores are the first source of inoculum infecting the primary host. They develop inside the host and produce thousands of Uredinia that break

through the plant epidermis. Stem rust Urediniospores are characterised by a brown-red colour. These last can infect adjacent plants and produce new Urediniospores, leading to exponential growth. This rapid and massive multiplication is the very reason of the disease severity and the fear that it can create. Furthermore, Urediniospores can be disseminated through wind over very long distances, causing epidemics on a scale as large as inter-continental; i.e. some strains in Australia are known to originate from Africa (Watson and De Sousa, 1983). When the crop season comes to an end and the wheat begins to ripen, black Teliospores are formed, from which Basidiospores are produced. Upon infection of the alternate host and conjugation of Pycniospores with receptive hyphae, Aeciospores are produced and the cycle starts again.

III.2. Control

Removing common barberry, especially around wheat fields, is a primary way to limit the development of stem rust. Important eradication campaigns were successfully undertaken in the first half of the 20th century in the US and still continue today. Without the alternate host, stem rust cannot complete its life cycle, hence removing a significant and early source of inoculum and reducing the genetic variation in the fungus population by eliminating the sexual cycle.

However, in regions with suitable conditions, Urediniospores can survive through the off-season on volunteer wheat and non-crop grass hosts, thus being able to start a new epidemic the year after. The airborne ability of Urediniospores is also an important factor for rust epidemics, especially in regions where the pathogen cannot survive all year round. Spores carried by the wind from warm regions become new inoculum in colder regions. The "*Puccinia* Pathway" is an example of this mode of dispersal, where the pathogen overwinters in southern USA and Mexico and is brought back further north through south-north winds (Stakman, 1957).

III.2.1. Genetic resistance

Because of its capability of dispersion, the most effective and economic way to control stem rust is the use of genetic resistance in wheat cultivars. Many resistance genes are already present in various hexaploid wheat varieties and even more in tetraploid wheat and related species, which are important sources of new rust resistance genes in hexaploid wheat.

Through breeding programs, these resistances can be incorporated to agronomically important cultivars. Resistance genes in wheat are generally categorized as race-specific (R-genes) or adult plant resistance (APR-genes).

R-genes confer resistance against some races of a pathogen but are ineffective against others, as determined by the gene-for-gene relationship between host resistance and pathogen avirulence genes (see "Plant immunity"). They usually provide effective resistance and involve hypersensitive reactions. They can be expressed in seedlings and adult plants. Only three R-genes conferring resistance to *Pgt* have been cloned so far: *Rpg1*, *Sr33* and *Sr35* (Brueggeman et al., 2002; Periyannan et al., 2013; Saintenac et al., 2013). The first was found in barley and encodes a protein kinase with two tandem kinase domains while the others encode classical NB-LRR proteins. Another gene in barley, *Rpg5*, which encodes an atypical resistance protein with NB and LRR domains as well as a kinase domain, alone provides resistance to an isolate of rye stem rust (*P. graminis* f. sp. *secalis*). This gene also confer resistance to multiple *Pgt*, including Ug99, when combined with several other genes, notably the wheat stem rust R-gene, *Rpg4* (Arora et al., 2013; Wang et al., 2013). These genes have not been introgressed to wheat.

The resistance provided by APR-genes is not as complete as that provided by R-genes. APR genes do not completely prevent pathogen growth; however, they significantly slow the rust growth (slow rusting) to a point where the plant can develop until harvest without serious damage. Contrary to R-genes, APR-genes are effective against a broad spectrum of races and even species. They also encode proteins that are usually distinct from those of R-genes. For instance, the APR-gene, *Lr34*, encodes an ABC-transporter that provides resistance against wheat leaf and stripe rust as well as powdery mildew (*Blumeria graminis*) (Krattinger et al., 2009). As suggested by their name, APR-genes usually provide resistance in plants at the reproductive stage and not seedlings. Aside from *Lr34,* which provides resistance to *Pgt* in some genetic backgrounds, no wheat APR-gene effective against *Pgt* has been cloned.

At least 50 stem rust resistance genes have been catalogued (Singh et al., 2012) and all except *Sr2* have the characteristics of R-genes. Most do not provide full resistance, but allow the development of small to medium Uredinia surrounded by chlorosis or necrosis. Some of these genes have been successfully used for years, but many were discarded, as they are no longer effective against pathogen populations in constant evolution. Although barberry eradication greatly diminishes the rust genetic diversity, new virulent races can appear

through mutation and genetic resistances in wheat can lose effectiveness after only a few years of widespread use.

To avoid the waste of valuable resistance genes through these "boom-and-bust" cycles, the best strategy appears to be to combine several of them (Singh et al., 2012). This approach, also called pyramiding, limits the outbreak of races virulent on one gene by controlling them with other genes. Chances of the pathogen to simultaneously evolve virulence to multiple major R-genes and/or minor APR-genes are then significantly lower. Djian-Caporalino et al. (2014) demonstrated that pyramiding two R-genes totally suppressed the emergence of virulent isolates of root-knot nematodes in pepper (*Capsicum annuum*), over a 3-years period, whereas conventional systems were not as efficient. This strategy, however, is only efficient if globally followed by all breeding programs. Indeed, stacking genes in one cultivar would turn out to be useless if the pathogen can adapt on nearby cultivars that only carry one of the genes. Good agricultural practices, such as the removal of common barberry, are also essential.

Because stem rust can adapt so quickly and develop new virulence, current resistance genes should not be expected to remain effective indefinitely. Even genes used for decades are likely to be overcome sooner or later. With the appearance of Ug99, *Sr31* is the perfect example.

III.3. Ug99

Sr31 has been transferred in wheat through the introgression of the chromosome fragment 1BL.1RS from rye (*Secale cereale*). This gene was so effective that it has been used, and still is, in many cultivars all over the world. After more than three decades of successful control, wheat stem rust was significantly on the decline until the emergence of Ug99, a highly virulent race against which *Sr31* is not effective. Ug99 is also referred as TTKSK using the North American nomenclature system (Jin et al., 2008; Roelfs, 1988). Ug99 was first observed in Uganda in 1999 (Pretorius et al., 2000). Since then, several variants of Ug99 appeared, such as TTKST and TTTSK, which are virulent to *Sr24* and *Sr36* respectively, both important resistance genes widely used and initially thought to be valuable sources of resistance against Ug99 (Jin et al., 2008, 2009). Currently, it is estimated that 80-90% of global wheat cultivars are susceptible to at least one of the suite of races of Ug99.

Ug99 and variants have now spread to Kenya and Ethiopia where they caused major epidemics (Wanyera et al., 2006), to Sudan, Yemen, Iran and South Africa (Nazari et al., 2009; Pretorius et al., 2010). It is predicted that these races will migrate to North Africa, the Middle East, Asia and beyond, menacing global food security. The threat posed by Ug99 is serious, and led in 2005 to the foundation of the "Borlaug Global Rust Initiative" (www.globalrust.org). This institute coordinates international efforts to fight Ug99 and facilitates collaborations between scientists, breeders and policymakers. Its main objective is to systematically reduce the world's vulnerability to wheat rusts.

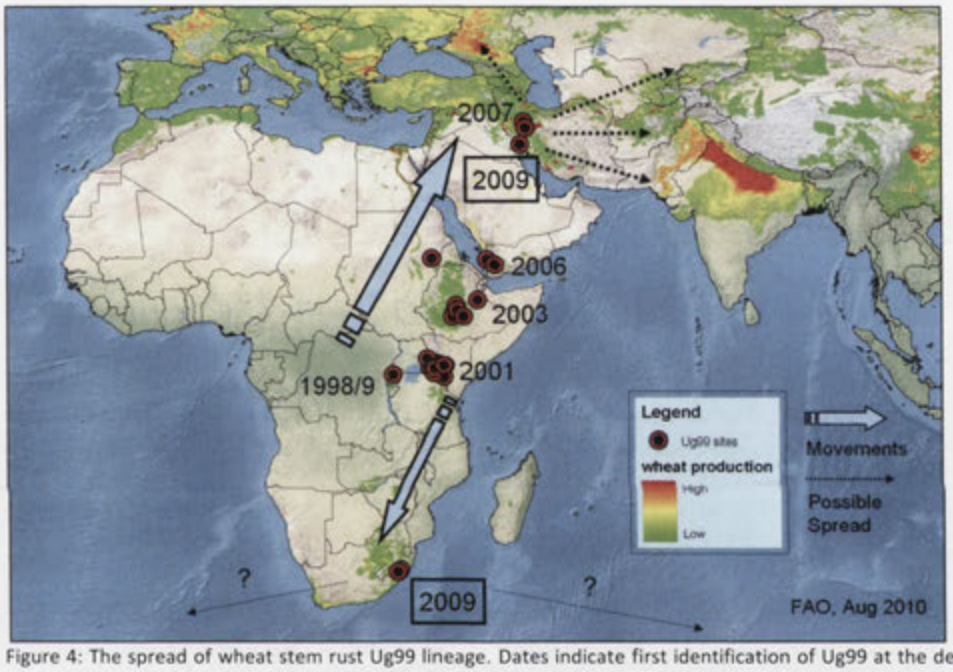


Figure 4: The spread of wheat stem rust Ug99 lineage. Dates indicate first identification of Ug99 at the designated sites. Source: FAO (http://www.fao.org/agriculture/crops/rust/stem/rust-report/stem-ug99racettksk/en/)

Replacing susceptible cultivars by durable resistant ones in regions struck by Ug99 is a priority for limiting its spread. Several resistant genes effective against Ug99 and variants are already known (Singh et al., 2012), but most of them only provide moderate resistance. In order to obtain significant resistance but also to avoid fast breakdown, these genes must be used in combination. Marker-assisted selection (MAS) is a powerful tool to achieve this quickly. By using molecular markers that are linked to genes controlling particular traits, MAS allows the selection of these genes without the necessity to phenotype the plants for the corresponding traits, which is time-consuming and sometimes complicated.

However, tightly linked markers for resistant genes are not always available as only very few resistance genes have been cloned. *Sr33* and *Sr35* are the only cloned resistance genes effective against Ug99 in wheat (Periyannan et al., 2013; Saintenac et al., 2013). Put together, these two genes provide resistance to all known *Pgt* races. Efforts are being made to combine these genes as well as others, such as *Sr2*, the only known APR-gene effective against stem rust, into elite cultivars for deployment in high-risk regions.

While this approach is expected to greatly reduce the threat posed by Ug99, the breakdown of *Sr31* demonstrated that there is no permanent solution when it comes to plant-pathogen interaction. This is especially true when good breeding and agricultural practices are not always respected, enhancing the risk of new outbreaks. Consequently, discovering new sources of resistance must remain a priority.

IV. Rust resistance suppression

While finding new genes of resistance is definitely one of the best strategies for fighting wheat rusts, resistance genes can be discarded, but not necessarily because of changes in pathogen virulence. Indeed, some genes can be very effective against rusts but are not used for various reasons. In some cases, these genes were introgressed in a particular wheat variety, along with other linked genes that confer undesirable traits. Due to their linkage with the gene of interest, these undesirable genes can be difficult to remove, especially when tightly linked markers are not available. This phenomenon is referred as linkage drag. Good markers remain the major tool for efficient breeding programs.

In other cases, the introgression of R genes into, for example, hexaploid wheat does not confer resistance to a wheat pathogen, although the gene was found effective against that pathogen in the variety or species from which it was sourced. This seems especially true with genes from wheat related species. Indeed, many studies showed that the resistance of genes from lower ploidy species was not expressed when transferred to hexaploid wheat. The difficulties encountered for transferring R genes appear to be of genetic nature and in many cases could be due to the presence of suppressors that inhibit the expression of resistance.

There are numerous examples of suppressors of rust resistance in wheat. Bai and Knott (1992) showed the presence of leaf and stem rust resistance suppressors on several D-genome chromosomes of common wheat. Innes and Kerber (1994) showed that leaf rust

resistance of *Ae. tauschii* was suppressed by loci on the A and B genomes when crossed with hexaploid wheat. In Knott et al. (2005), leaf and stem rust resistance genes found in Emmer tetraploid wheat were, after crossing with bread wheat, suppressed by genes also found on the A and B genomes. There are many cases of synthetic hexaploid wheats that do not express resistances present in the parental lines (Assefa and Fehrmann, 2000, 2004; Chen et al., 2013; Kema et al., 1995; Ma et al., 1995).

In contrast, there are few examples of progenies gaining resistance that were not expressed in the parental lines (Knott, 2000a, 2000b). *SuLr23*, a suppressor of *Lr23* which is on 2BS, was mapped on the homologous region of 2DS (Nelson et al., 1997). Suppression of resistance is not only restricted to rusts. Recently, *Pm8*, which confers resistance to powdery mildew and was introgressed in wheat from rye, was found suppressed by the orthologous wheat gene *Pm3* (Hurni et al., 2013; McIntosh et al., 2011). The molecular mechanism of suppression remains yet to be elucidated, but the cloning of both genes should give insight. However, this is not the case for most suppressors. Finding them is, therefore, the first step in order to understand their actual role, as resistance suppression appears to be a negative trait. The high frequency of suppressors in wheat suggests that they may confer a selective advantage. For example, they may neutralize fitness penalties caused by resistance genes in the absence of pathogens.

Nevertheless, characterising these suppressors is essential in order to prevent their transfer in breeding programs and to reveal masked resistance genes. The Canthatch 7DL suppressor (7DL-Sup) appears of particular interest as the genes it inhibits were found to confer robust resistance against Ug99 (Lagudah and Jin, unpublished).

IV.1. Canthatch 7DL suppressor

More than 30 years ago, (Kerber and Green, 1980) observed that the hexaploid wheat cultivar Canthatch (CTH; AABBDD, 2n=42) was susceptible to several stem rust races, whereas the tetraploid derived from CTH (AABB, 2n=28), which does not possess the D genome, was resistant. A comparable resistance was also observed in CTH nullisomic 7D (CTH-N7D) and CTH ditelosomic 7DS (CTH-Dt7DS), which respectively lack the whole chromosome 7D or only the long arm (Kerber, 1991; Kerber and Green, 1980). In contrast, the reaction of CTH ditelosomic 7DL (CTH-Dt7DL) was similar to that of CTH. This indicates that a gene(s) located in CTH on the

long arm of chromosome 7D suppresses resistance conferred by other genes, most likely on the A and B genomes.

CTH was developed by backcrossing six times Kenya Farmer into the recurrent parent Thatcher (Tc), the probable donor of the 7DL-Sup (Kerber, 1991). The origin of the suppressor was further investigated by replacing the chromosome 7DL of CTH by the one from six common wheat cultivars, including Chinese Spring (CS), as well as six synthetic hexaploid wheat, of which the D genome is from different strains of *Ae. tauschii* (Kerber, 1983). All F_1 progenies were as susceptible as CTH indicating that the 7DL-Sup may be prevalent in common wheat cultivars and *Ae. tauschii*, and probably originates from this last (Kerber, 1983).

Two sets of non-suppressor (NS) mutants were developed from CTH using ethyl methanesulphonate (EMS). The first set (Kerber, 1991), developed from CTH-K, a particular selection of CTH with a fixed genotype, comprises two mutants, CTH-NS1 and CTH-NS2, and the second set (Williams et al., 1992) comprises fifteen mutants, named from North Dakota Stem rust Resistant (NDSR)-1 to 15, of which NDSR12 was ditelosomic 7DS. All mutants were resistant to the stem rust races to which CTH-K was susceptible. Genetic studies showed that the resistance was monogenic and indicated that a recessive, non-suppressing mutation had been induced at or near the suppressor locus on 7DL. Further experiments confirmed that the mutation in CTH-NS behaved as a non-suppressing allele (Kerber and Aung, 1995). The non-suppressor locus was found to be independent of the centromere, but was not mapped beyond this (Kerber, 1991).

Crosses between the NDSR mutants and CS monosomic 7DL indicated that up to three recessive resistance genes in CTH may be suppressed by the 7DL-Sup (Williams et al., 1992). CTH is known for carrying the stem rust resistance genes *Sr5*, *Sr7a*, *Sr9g*, *Sr12* and *Sr16* but these genes are not likely to be the ones that are suppressed as they confer resistance to particular stem rust races (Kerber and Green, 1980). Because CTH is a derivative of Tc, from which all the above genes except *Sr7a* and probably the suppressor originated, other genes found in Tc may also be present in CTH, including those that are suppressed by the suppressor. Knott (2009) found that Tc may also carry Sr6, Sr7a, Sr8a and Sr9d as well as several other unidentified genes and suppressors.

As previously mentioned, the NS mutation confers resistance to Ug99 and derived races TTKST and TTTSK (Lagudah and Jin, unpublished). The genes inhibited by the 7DL-Sup are

inferred to be responsible for this resistance. This makes the study of this suppressor very attractive in order to unmasked new resistance genes. All the more so as this suppressor seems to be quite common in wheat (Kerber, 1983) and would first need to be inactivated in order to use the corresponding suppressed genes. Also, contrary to any other suspected suppressors, NS mutants are already available, which can greatly help for characterising it.

V. Project outlines

The initial objectives of this project were to fine map the 7DL-Sup and characterise its mode of action. To do this, a $F_{2:3}$ mapping population descending from the cross between Columbus (Col) and Columbus-NS766 (Col-NS766) was developed. Like CTH, Col is a Tc backcross-derived cultivar and was thought to possess the 7DL-Sup. Col-NS766 was developed by backcrossing CTH-NS into Col in order to introgress the NS allele into the Col background (Kerber, 1991). These two lines were chosen for mapping because they were thought to carry either the 7DL-Sup or the NS allele, and were genetically similar while having a good level of polymorphism between each other, essential for mapping studies, which the CTH lines lack. However, in the course of the genetic mapping, it became clear that the locus being mapped was not located on the chromosome 7DL but on 7AL.

Because of the extensive genetic evidence for the existence of the suppressor on chromosome 7DL, this result was surprising and implied that the mapping population was not segregating for the 7DL-Sup but for a locus of resistance on 7AL. The fact that this resistance was expressed in CTH-NS but not in CTH, indicated that the 7AL locus is normally suppressed by the 7DL-Sup. The finding that introgression of the 7AL locus into Col produced resistance, suggested that, contrary to expectation, Col did not in fact possess the 7DL-Sup or the suppressed genes. Thus, only the introgression of a resistance locus would confer resistance. An independent backcrossed line similar to Col-NS766, Col-NS765, which also had the 7AL locus introgressed, corroborated this hypothesis. Although this discovery compromised the initial objectives of the project, characterising the 7AL locus remained as an important goal, since the locus conferred resistance to Ug99.

Genetic mapping of the 7AL was done by phenotyping F_3 families using two stem rust races, including Ug99, and by genotyping F_2 plants using several sets of molecular markers. The first two sets of markers were reported in the literature to be located on chromosome 7D, and

are detailed in chapter 1. The other two sets were discovered using recently developed genotyping technologies: i) an Infinium BeadChip comprising nearly 9,000 wheat single nucleotide polymorphisms (SNPs) and ii) genotyping by sequencing (GBS). These technologies are detailed in chapter 2. As part of the characterisation of the resistance response, the rust development in susceptible and resistant lines was described through histological observations. Whole transcriptome analysis was also undertaken using RNA-Sequencing (RNA-Seq) and is detailed in chapter 3.

Project contributions

Several robust markers tightly linked to the 7AL resistance locus were developed. These markers could help the transfer of this locus into agronomically important wheat varieties via MAS, especially in regions struck by Ug99. Moreover, the discovery of several genes potentially involved in the resistance response gives some insights into the mechanism of resistance and may help in the development of new strategies for fighting rust.

REFERENCES

Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature *408*, 796–815.

Arora, D., Gross, T., and Brueggeman, R. (2013). Allele Characterization of Genes Required for rpg4-Mediated Wheat Stem Rust Resistance Identifies *Rpg5* as the R Gene. Phytopathology *103*, 1153–1161.

Assefa, S., and Fehrmann, H. (2000). Resistance to wheat leaf rust in *Aegilops tauschii* Coss. and inheritance of resistance in hexaploid wheat. Genet. Resour. Crop Evol. *47*, 135–140.

Assefa, S., and Fehrmann, H. (2004). Evaluation of *Aegilops tauschii* Coss. for resistance to wheat stem rust and inheritance of resistance genes in hexaploid wheat. Genet. Resour. Crop Evol. *51*, 663–669.

Ausubel, F.M. (2005). Are innate immune signaling pathways in plants and animals conserved? Nat. Immunol. *6*, 973–979.

Ayliffe, M., Devilla, R., Mago, R., White, R., Talbot, M., Pryor, A., and Leung, H. (2011). Nonhost resistance of rice to rust pathogens. Mol. Plant. Microbe Interact. *24*, 1143–1155.

Bai, D., and Knott, D.R. (1992). Suppression of rust resistance in bread wheat (*Triticum aestivum* L.) by D-genome chromosomes. Genome *35*, 276–282.

Brenchley, R., Spannagl, M., Pfeifer, M., Barker, G.L.A., D'Amore, R., Allen, A.M., McKenzie, N., Kramer, M., Kerhornou, A., Bolser, D., et al. (2012). Analysis of the bread wheat genome using whole-genome shotgun sequencing. Nature *491*, 705–710.

Brueggeman, R., Rostoks, N., Kudrna, D., Kilian, A., Han, F., Chen, J., Druka, A., Steffenson, B., and Kleinhofs, A. (2002). The barley stem rust-resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases. Proc. Natl. Acad. Sci. *99*, 9328–9333.

Chandra, S., and Low, P.S. (1995). Role of phosphorylation in elicitation of the oxidative burst in cultured soybean cells. Proc. Natl. Acad. Sci. *92*, 4120–4123.

Chapman, V., Miller, T.E., and Riley, R. (1976). Equivalence of the A genome of bread wheat and that of *Triticum urartu*. Genet. Res. *27*, 69–76.

Chen, W., Liu, T., and Gao, L. (2013). Suppression of stripe rust and leaf rust resistances in interspecific crosses of wheat. Euphytica *192*, 339–346.

Derksen, H., Rampitsch, C., and Daayf, F. (2013). Signaling cross-talk in plant disease resistance. Plant Sci. *207*, 79–87.

Djian-Caporalino, C., Palloix, A., Fazari, A., Marteu, N., Barbary, A., Abad, P., Sage-Palloix, A.-M., Mateille, T., Risso, S., and Lanza, R. (2014). Pyramiding, alternating or mixing: comparative performances of deployment strategies of nematode resistance genes to promote plant resistance efficiency and durability. BMC Plant Biol. *14*, 53.

Dodds, P.N., and Rathjen, J.P. (2010). Plant immunity: towards an integrated view of plant–pathogen interactions. Nat. Rev. Genet. *11*, 539–548.

Dvorak, J. (1976). The relationship between the genome of *triticum urartu* and the A and B genomes of *triticum aestivum*. Can. J. Genet. Cytol. *18*, 371–377.

Dvorak, J., Terlizzi, P. di, Zhang, H.-B., and Resta, P. (1993). The evolution of polyploid wheats: Identification of the A genome donor species. Genome *36*, 21–31.

FAO (2013). The state of food insecurity in the world.

Felix, G., Regenass, M., and Boller, T. (1993). Specific perception of subnanomolar concentrations of chitin fragments by tomato cells: induction of extracellular alkalinization, changes in protein phosphorylation, and establishment of a refractory state. Plant J. *4*, 307–316.

Felix, G., Duran, J.D., Volko, S., and Boller, T. (1999). Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. Plant J. *18*, 265–276.

Flor, H.H. (1971). Current status of the gene-for-gene concept. Annu. Rev. Phytopathol. *9*, 275–296.

Fu, Z.Q., and Dong, X. (2013). Systemic Acquired Resistance: turning local infection into global defense. Annu. Rev. Plant Biol. *64*, 839–863.

Gill, B.S., Li, W., Sood, S., Kuraparthy, V., Friebe, B.R., Simons, K.J., Zhang, Z., and Faris, J.D. (2007). Genetics and genomics of wheat domestication-driven evolution. Isr. J. Plant Sci. *55*, 223–229.

Göhre, V., Spallek, T., Häweker, H., Mersmann, S., Mentzel, T., Boller, T., de Torres, M., Mansfield, J.W., and Robatzek, S. (2008). Plant pattern-recognition receptor FLS2 is directed for degradation by the bacterial ubiquitin ligase AvrPtoB. Curr. Biol. *18*, 1824–1832.

Göhre, V., Jones, A.M., Sklenár, J., Robatzek, S., and Weber, A.P. (2012). Molecular crosstalk between PAMP-triggered immunity and photosynthesis. Mol. Plant. Microbe Interact. *25*, 1083–1092.

Granado, J., Felix, G., and Boller, T. (1995). Perception of fungal sterols in plants (subnanomolar concentrations of ergosterol elicit extracellular alkalinization in tomato cells). Plant Physiol. *107*, 485–490.

Gust, A.A., Biswas, R., Lenz, H.D., Rauhut, T., Ranf, S., Kemmerling, B., Gotz, F., Glawischnig, E., Lee, J., Felix, G., et al. (2007). Bacteria-derived peptidoglycans constitute pathogen-associated molecular patterns triggering innate immunity in *Arabidopsis*. J. Biol. Chem. *282*, 32338–32348.

Haider, N. (2013). The origin of the B-genome of bread wheat (*Triticum aestivum* L.). Russ. J. Genet. *49*, 263–274.

Heun, M. (1997). Site of Einkorn Wheat Domestication Identified by DNA Fingerprinting. Science *278*, 1312–1314.

Hogenhout, S.A., Van der Hoorn, R.A., Terauchi, R., and Kamoun, S. (2009). Emerging concepts in effector biology of plant-associated organisms. Mol. Plant. Microbe Interact. *22*, 115–122.

Hurni, S., Brunner, S., Buchmann, G., Herren, G., Jordan, T., Krukowski, P., Wicker, T., Yahiaoui, N., Mago, R., and Keller, B. (2013). Rye *Pm8* and wheat *Pm3* are orthologous genes and show evolutionary conservation of resistance function against powdery mildew. Plant J. *76*, 957–969.

Innes, R.L., and Kerber, E.R. (1994). Resistance to wheat leaf rust and stem rust in *Triticum tauschii* and inheritance in hexaploid wheat of resistance transferred from *T. tauschii*. Genome *37*, 813–822.

International Rice Genome Sequencing Project (2005). The map-based sequence of the rice genome. Nature *436*, 793–800.

Jia, J., Zhao, S., Kong, X., Li, Y., Zhao, G., He, W., Appels, R., Pfeifer, M., Tao, Y., Zhang, X., et al. (2013). *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. Nature *496*, 91–95.

Jin, Y., Szabo, L.J., Pretorius, Z.A., Singh, R.P., Ward, R., and Fetch Jr, T. (2008). Detection of virulence to resistance gene *Sr24* within race TTKS of *Puccinia graminis* f. sp. *tritici*. Plant Dis. *92*, 923–926.

Jin, Y., Szabo, L.J., Rouse, M.N., Fetch Jr, T., Pretorius, Z.A., Wanyera, R., and Njau, P. (2009). Detection of virulence to resistance gene *Sr36* within the TTKS race lineage of *Puccinia graminis* f. sp. *tritici*. Plant Dis. *93*, 367–370.

Kema, G.H.J., Lange, W., and Van Silfhout, C.H. (1995). Differential suppression of stripe rust resistance in synthetic wheat hexaploids derived from *Triticum turgidum* subsp. *dicoccoides* and *Aegilops squarrosa*. Phytopathology *85*, 425–429.

Kerber, E.R. (1983). Suppression of Rust Resistance in Amphiploids of *Triticum*. In Proc. 6th International Wheat Genetics Symposium, (Kyoto, Japan), pp. 813–817.

Kerber, E.R. (1991). Stem-rust resistance in "Canthatch" hexaploid wheat induced by a nonsuppressor mutation on chromosome 7DL. Genome *34*, 935–939.

Kerber, E.R., and Aung, T. (1995). Confirmation of nonsuppressor mutation of stem rust resistance in "Canthatch" common wheat. Crop Sci. *35*, 743–744.

Kerber, E.R., and Green, G.J. (1980). Suppression of stem rust resistance in the hexaploid wheat cv. Canthatch by chromosome 7DL. Can. J. Bot. *58*, 1347–1350.

Knott, D.R. (2000a). Inheritance of resistance to stem rust in Medea durum wheat and the role of suppressors. Crop Sci. *40*, 98–102.

Knott, D.R. (2000b). Inheritance of resistance to stem rust in "Triumph 64"winter wheat. Crop Sci. *40*, 1237–1241.

Knott, D.R. (2009). Genes for stem rust resistance in Thatcher wheat. Can. J. Plant Sci. *89*, 1003–1008.

Knott, D.R., Bai, D., and Zale, J. (2005). The transfer of leaf and stem rust resistance from wild emmer wheats to durum and common wheat. Can. J. Plant Sci. *85*, 49–57.

Krattinger, S.G., Lagudah, E.S., Spielmeyer, W., Singh, R.P., Huerta-Espino, J., McFadden, H., Bossolini, E., Selter, L.L., and Keller, B. (2009). A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. Science *323*, 1360–1363.

Ling, H.-Q., Zhao, S., Liu, D., Wang, J., Sun, H., Zhang, C., Fan, H., Li, D., Dong, L., Tao, Y., et al. (2013). Draft genome of the wheat A-genome progenitor *Triticum urartu*. Nature *496*, 87–90.

Ma, H., Singh, R.P., and Mujeeb-Kazi, A. (1995). Suppression/expression of resistance to stripe rust in synthetic hexaploid wheat (*Triticum turgidum* x *T. tauschii*). Euphytica *83*, 87–93.

Marone, D., Russo, M., Laidò, G., De Leonardis, A., and Mastrangelo, A. (2013). Plant nucleotide binding site–leucine-rich repeat (NBS-LRR) genes: active guardians in host defense responses. Int. J. Mol. Sci. *14*, 7302–7326.

Matsuoka, Y. (2011). Evolution of polyploid *Triticum* Wheats under cultivation: the role of domestication, natural hybridization and allopolyploid speciation in their diversification. Plant Cell Physiol. *52*, 750.

McFadden, E.S., and Sears, E.R. (1946). The origin of *Triticum spelta* and its free-threshing hexaploid relatives. J. Hered. *37*, 107–116.

McIntosh, R.A., Zhang, P., Cowger, C., Parks, R., Lagudah, E.S., and Hoxha, S. (2011). Rye-derived powdery mildew resistance gene *Pm8* in wheat is suppressed by the *Pm3* locus. Theor. Appl. Genet. *123*, 359–367.

Melotto, M., Underwood, W., Koczan, J., Nomura, K., and He, S.Y. (2006). Plant stomata function in innate immunity against bacterial invasion. Cell *126*, 969–980.

Monaghan, J., and Zipfel, C. (2012). Plant pattern recognition receptor complexes at the plasma membrane. Curr. Opin. Plant Biol. *15*, 349–357.

Mur, L.A.J., Kenton, P., Lloyd, A.J., Ougham, H., and Prats, E. (2008). The hypersensitive response; the centenary is upon us but how much do we know? J. Exp. Bot. *59*, 501–520.

Murray, G.M., and Brennan, J.P. (2009). Estimating disease losses to the Australian wheat industry. Australas. Plant Pathol. *38*, 558.

Muthamilarasan, M., and Prasad, M. (2013). Plant innate immunity: An updated insight into defense mechanism. J. Biosci. *38*, 433–449.

Nazari, K., Mafi, M., Yahyaoui, A., Singh, R.P., and Park, R.F. (2009). Detection of wheat stem rust (*Puccinia graminis* f. sp. *tritici*) race TTKSK (Ug99) in Iran. Plant Dis. *93*, 317–317.

Nelson, J.C., Singh, R.P., Autrique, J.E., and Sorrells, M.E. (1997). Mapping genes conferring and suppressing leaf rust resistance in wheat. Crop Sci. *37*, 1928–1935.

Periyannan, S., Moore, J., Ayliffe, M., Bansal, U., Wang, X., Huang, L., Deal, K., Luo, M., Kong, X., Bariana, H., et al. (2013). The gene *Sr33*, an ortholog of barley *Mla* genes, encodes resistance to wheat stem rust race ug99. Science *341*, 786–788.

Pretorius, Z.A., Singh, R.P., Wagoire, W.W., and Payne, T.S. (2000). Detection of virulence to wheat stem rust resistance gene *Sr31* in *Puccinia graminis* . f. sp. *tritici* in Uganda. Plant Dis. *84*, 203–203.

Pretorius, Z.A., Bender, C.M., Visser, B., and Terefe, T. (2010). First report of a *Puccinia graminis* f. sp. *tritici* race virulent to the *Sr24* and *Sr31* wheat stem rust resistance genes in South Africa. Plant Dis. *94*, 784–784.

Reina-Pinto, J.J., and Yephremov, A. (2009). Surface lipids and plant defenses. Plant Physiol. Biochem. *47*, 540–549.

Rodriguez-Galvez, E., and Mendgen, K. (1995). Cell wall synthesis in cotton roots after infection with Fusarium oxysporum. Planta *197*, 535–545.

Roelfs, A.P. (1988). An international system of nomenclature for *Puccinia graminis* f. sp. *tritici*. Phytopathology *78*, 526.

Roelfs, A.P., Singh, R.P., Saari, E.E., and International Maize and Wheat Improvement Center (1992). Rust diseases of wheat: concepts and methods of disease management (Mexico, D.F: CIMMYT).

Ronald, P.C., and Beutler, B. (2010). Plant and animal sensors of conserved microbial signatures. Science *330*, 1061–1064.

Saintenac, C., Zhang, W., Salcedo, A., Rouse, M.N., Trick, H.N., Akhunov, E., and Dubcovsky, J. (2013). Identification of wheat gene *Sr35* that confers resistance to ug99 stem rust race group. Science *341*, 783–786.

Shan, L., He, P., Li, J., Heese, A., Peck, S.C., Nürnberger, T., Martin, G.B., and Sheen, J. (2008). Bacterial effectors target the common signaling partner BAK1 to disrupt multiple MAMP receptor-signaling complexes and impede plant immunity. Cell Host Microbe *4*, 17–27.

Shewry, P.R. (2009). Wheat. J. Exp. Bot. *60*, 1537–1553.

Singh, S., Singh, R.P., and Huerta-Espino, J. (2012). Stem rust. In disease resistance in wheat, (Wallingford, Oxfordshire; Cambridge, MA: CABI), pp. 18 – 32.

Smith, D.B., and Flavell, R.B. (1975). Characterisation of the wheat genome by renaturation kinetics. Chromosoma *50*, 223–242.

Spoel, S.H., and Dong, X. (2012). How do plants achieve immunity? Defence without specialized immune cells. Nat. Rev. Immunol. *12*, 89–100.

Stakman, E.C. (1957). Problems in preventing plant disease epidemics. Am. J. Bot. *44*, 259–267.

Wang, X., Richards, J., Gross, T., Druka, A., Kleinhofs, A., Steffenson, B., Acevedo, M., and Brueggeman, R. (2013). The rpg4 -mediated resistance to wheat stem rust (*Puccinia graminis*) in barley (*Hordeum vulgare*) requires *Rpg5*, a second NBS-LRR gene, and an actin depolymerization factor. Mol. Plant. Microbe Interact. *26*, 407–418.

Wanyera, R., Kinyua, M.G., Jin, Y., and Singh, R.P. (2006). The spread of stem rust caused by *Puccinia graminis* f. sp. *tritici*, with virulence on *Sr31* in wheat in eastern africa. Plant Dis. *90*, 113–113.

Watson, I.A., and De Sousa, C.N.A. (1983). Long distance transport of spores of *Puccinia graminis tritici* in the southern hemisphere. Proc Linn Soc NSW *106*.

Williams, N.D., Miller, J.D., and Klindworth, D.L. (1992). Induced mutations of a genetic suppressor of resistance to wheat stem rust. Crop Sci. *32*, 612–616.

Chapter 2

Initial attempt to map the Canthatch 7DL suppressor locus

ABSTRACT

Initial mapping of the chromosome 7DL suppressor of stem rust resistance was conducted using a $F_{2:3}$ mapping population derived from the cross between the susceptible Columbus and the resistant Columbus-NS766, alleged to carry the suppressor and the non-suppressor variants, respectively. Plants were phenotyped with two stem rust races, one of Australian origin, #313, and Ug99 (TTKSK). The two races gave similar results suggesting a common resistance mechanism. Chi-square tests indicated one-gene segregation. 27 SSR and 23 SNP markers found on chromosome 7DL were retrieved from the literature and tested on the mapping population. 2 SSRs (wmc0273 and cfa2040) were found to be linked to the resistance.

List of major abbreviations

7DL-Sup Canthatch 7DL suppressor; **Col** Columbus; **Col-NS** Columbus-NonSuppressor; **CTH** Canthatch; **CTH-NS** Canthatch-NonSuppressor; **DFA** DNA fragment analyzer; **IT** infection type; **NDSR** North Dakota Stem rust Resistant; **NS** non-suppressor; **Pgt** *Puccinia graminis* f. sp. *tritici*; **SNPs** single nucleotide polymorphisms; **SSR** simple sequence repeats; **Tc** Thatcher.

INTRODUCTION

One of the most successful approaches used in wheat to clone genes is map-based cloning (Krattinger et al., 2009). A good example is the cloning of the leaf rust resistance gene *Lr10* (Feuillet et al., 2003; Stein et al., 2000; Wicker et al., 2001). This method, illustrated in Figure 1, consists in identifying a candidate region, by using a mapping population derived from a cross between two parents, differing in the trait of interest (e.g., Resistance/Susceptibility). To do so, each individual of the progeny is phenotyped for the trait of interest and genotyped with molecular markers, in order to build a genetic map. The relationship between phenotypes and genotypes indicates the relative position of the gene of interest and the closest markers flanking the gene define the candidate region. Those are then used as starting points for the physical mapping of this region, which usually consists in isolating and sequencing DNA fragments from genomic libraries, such as bacterial artificial chromosomes (BACs). By chromosome walking, more BACs are then selected until no gap between the flanking markers remains. Finally, candidate genes are selected for functional analysis in order to uncover the gene conferring the trait.

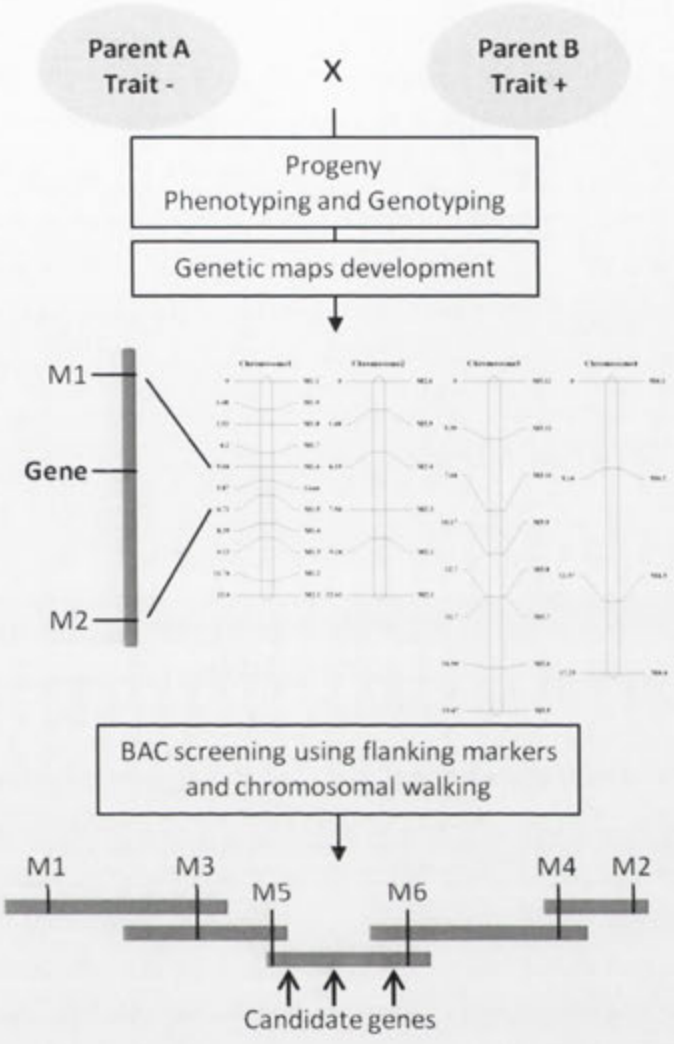


Figure 1: Map-based cloning strategy. The progeny descended from the cross between two parental lines differing in the trait of interest (+/-) is phenotyped for the trait and genotyped with molecular markers. Genetic maps are then developed and markers (M1/M2) flanking the gene conferring the trait are selected for the first round of BAC screening. From the selected BACs, new markers are used for additional screening and so on until the gap between both ends is closed. Following sequencing, candidate genes are finally selected for functional analysis.

Map-based cloning appears to be the strategy of choice for identifying the Canthatch 7DL suppressor (7DL-Sup). To date, too little is known about the 7DL-Sup to allow its rapid identification. Using monosomic and ditelosomic lines, the 7DL-Sup was found to be located on chromosome 7DL and independent of the centromere (Kerber, 1991). However, no linked markers are available. Progress with map-based cloning can be accelerated if markers, closely linked to the gene of interest, are known. In the absence of a well defined genetic interval that spans the 7DL-Sup locus, an attempt was made to develop a low resolution genetic mapping family, as a starting point to find markers linked to the locus.

Mapping population

Map-based cloning requires a population derived from lines which differ in the trait of interest, in this case, the suppression of resistance. One line must, therefore, carry the 7DL-Sup, whereas the other either does not have it or carries a non-functional allele. In addition, they must both possess the resistance gene(s) usually inhibited by the suppressor. As described in the introduction, several Canthatch (CTH) non-suppressor (NS) mutants are already available (Kerber, 1991; Williams et al., 1992). The use of mutants is preferable, as they can, upon sequencing of both parents, facilitate the selection of candidate genes. However, crossing CTH with one of the mutants was not found to be judicious, as these lines are completely monomorphic, except for few induced mutations. A higher level of polymorphism is necessary for marker development, hence lines must be divergent. Nonetheless, their genetic background should be similar in order to be phenotypically comparable.

Two lines seemed to display these requirements: Columbus (Col) and Columbus-NS766 (Col-NS766). Col is closely related to CTH, as it is also a Thatcher (Tc) backcross-derived cultivar. The full pedigree of those lines can be seen in Table 1. Col is susceptible to the races for which CTH and Tc are susceptible, thus, Col was suspected to also have the 7DL-Sup. Crosses between Col and the resistant CTH mutants, CTH-NS1 and CTH-NS2, provided clues for this assumption, as resistance was found in the progeny, presumably due to introgression of the NS locus in Col (Kerber, 1991). Two resistant lines were developed, by backcrossing CTH-NS1 and CTH-NS2, with Col as the recurrent parent, and selecting for resistance. The resulting BC_5F_4 lines, Col-NS765 and Col-NS766, possess most of the Col genetic background with, presumably, the NS locus from CTH-NS.

Line	Pedigree
Thatcher	MARQUIS/(TR.DR)IUMILLO//(HN-3001)MARQUIS/KANRED
Canthatch	THATCHER*6/KENYA-FARMER
Columbus	NEEPAWA*6/RL-4137
Neepawa	THATCHER*7/FRONTANA//THATCHER*6/KENYA-FARMER/3/THATCHER*2//FRONTANA/THATCHER
RL-4137	RL-2520//THATCHER*6/KENYA-FARMER

Table 1: Thatcher and Thatcher-derived lines pedigree.

To sum up, Col and Col-NS766 are closely related to CTH and Tc. They are presumed to carry the 7DL-Sup and the NS allele, respectively, as well as the corresponding resistance genes. Moreover, they are genetically similar with a good level of polymorphism at the introgressed locus. Thus, Col and Col-NS766 were chosen as the parental lines of choice for

map-based cloning of the 7DL-Sup. They were preferred to the original CTH lines because these last lack polymorphism, which is essential for mapping studies.

Col and Col-NS766 were crossed together and the F_2-derived F_3 ($F_{2:3}$) progeny were used to generate a low-resolution mapping population (Figure 2). This type of mapping population comprising two generations is more powerful than a single-generation one, as phenotyping multiple plants in F_3 is more accurate than phenotyping single plants in F_2. In addition, it allows the recognition of heterozygote F_2 by segregation in F_3 (Zhang and Xu, 2004). F_3 families were phenotyped for resistance to stem rust, while F_2 plants were genotyped with molecular markers. The $F_{2:3}$ progeny were descended from 5 F_1 plants (named A to E) and could be divided into two groups: (i) a first group developed at CSIRO Plant Industry, Canberra, before the start of the described work, with a total of 128 $F_{2:3}$ families from A and B F_1 plants (62 and 66 families, respectively), and (ii) a second group developed at the end of the first year of this project, with a total of 576 $F_{2:3}$ families from C, D and E F_1 plants (250, 212 and 114 families, respectively). The first group was predominantly used for mapping, due to its early availability and the greater number of phenotyping repetitions. The second group was developed with the aim of producing a higher resolution map of the 7DL-Sup, as a prelude to map-based cloning of the gene.

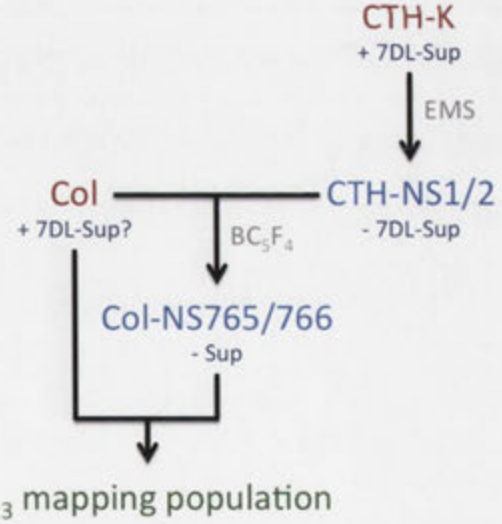


Figure 2: Development of the mapping population. CTH-K, which possesses the 7DL-Sup, was first mutated with EMS, giving the CTH-NS mutants, which possess the NS allele. These last were then backcrossed with the recurrent parent Col, which was thought to possess the 7DL-Sup, obtaining the Col-NS lines without the 7DL-Sup. The $F_{2:3}$ mapping population was then developed from the cross Col x Col-NS766.

Molecular markers

Until recently, before advancements in new technologies (discussed in the next chapter), developing molecular markers was considered difficult and time consuming. This was especially true in wheat and other polyploid species, for which homoeologous chromosomal regions, harbouring highly similar DNA sequences, can affect marker specificity. Thus, genetic mapping often relied on using previously characterised molecular markers reported in the literature and, often, with varying degrees of success. Although thousands of markers are available, they are based on various genotypes and are not necessarily polymorphic or easy to use for a particular mapping population. On the other hand, existing maps that are corroborated in consensus genetic maps, serve as a useful starting point when enrichment of markers, for a specified chromosomal region, is being sought. Two types of markers were used in this study; simple sequence repeats (SSR) and single nucleotide polymorphism (SNP) markers.

SSR markers, also known as microsatellites, are a series of bi-, tri-, tetra-nucleotide repeats, of which the length can vary between alleles (Semagn et al., 2006). Microsatellites are one of the most widely used markers in plants. They have the advantage of being highly polymorphic, even between related lines, and are co-dominant, which allows the detection of several alleles. Moreover, these markers are especially appropriate in wheat, as they can also differentiate between homoeologous regions, provided they are polymorphic. Technically, they only require a PCR and an agarose gel, though a DNA fragment analyzer (DFA) might be necessary to resolve small length differences. Figure 3 shows an example of results obtained from Col and Col-NS766 with a polymorphic SSR, using DFA.

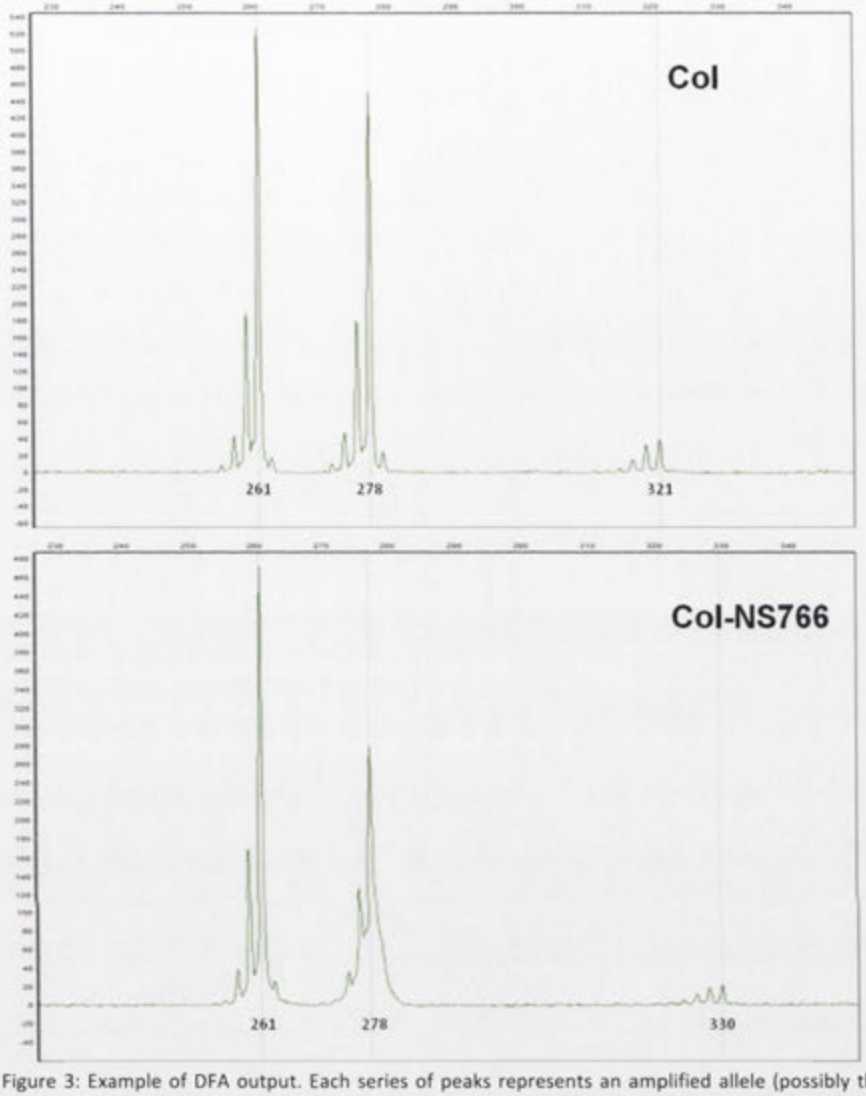


Figure 3: Example of DFA output. Each series of peaks represents an amplified allele (possibly the homoeologues regions in A, B and D genomes). Numbers represent their relative size. 3 fragments are present on Col and Col-NS766 but only one is polymorphic (321/330).

The second set of markers retrieved from the literature was based on SNPs. As suggested by the name, they are based on single-nucleotide differences. They are highly abundant and distributed throughout the genome (Semagn et al., 2006) and, as sequencing technologies improve, they are becoming more and more available.

In this chapter, an attempt of coarse mapping of the 7DL-Sup is described. Phenotypes were determined using two distinct races of stem rust and genotypes were determined using molecular markers, known in the literature to be located on chromosome 7DL.

RESULTS

I. Phenotyping

F_3 plants from the cross Col x Col-NS766 were phenotyped for resistance against wheat stem rust (*Puccinia graminis* f. sp. *tritici* abbreviated as *Pgt*) using two distinct races: (i) race #313, tested at the University of Sydney, Australia and (ii) Ug99 (TTKSK), tested at the USDA-ARS, Minnesota, USA. Infection type (IT) was assessed using the stem rust scoring scale (McIntosh et al., 1995).

Infection Type	Host Response	Symptoms
0	Immune	No visible uredia
;	Very resistant	Hypersensitive flecks
1	Resistant	Small uredia with necrosis
2	Resistant to moderately resistant	Small to medium sized uredia with green islands and surrounded by necrosis or chlorosis
3	Moderately resistant/moderately susceptible	Medium sized uredia with or without chlorosis
4	Susceptible	Large uredia without chlorosis
X	Resistant	Heterogeneous, similarly distributed over the leaves
Y	?	Variable size with larger uredia towards the tip
Z	?	Variable size with larger uredia towards the leaf base

Table 2: Major Infection type classes for stem rust and leaf rust (McIntosh et al., 1995)

I.1. Race #313

Because the use of Ug99 is prohibited in Australia, a surrogate stem rust strain was needed. A number of Australian stem rust races were tested but most were found avirulent on Col, implying that this line carries several resistance genes (pers. com. Robert McIntosh). Race #313 (pathotype 34-1,2,3,5,6,7) was the only one permitting good discrimination between Col and Col-NS766. Most experiments in this project were based on this race. However, the use of #313 is restricted. For this reason, seedlings were inoculated by collaborators at the Plant Breeding Institute (PBI), at the University of Sydney, in Cobbitty. Disease phenotypes were scored on the first leaf, approximately 2 weeks after inoculation.

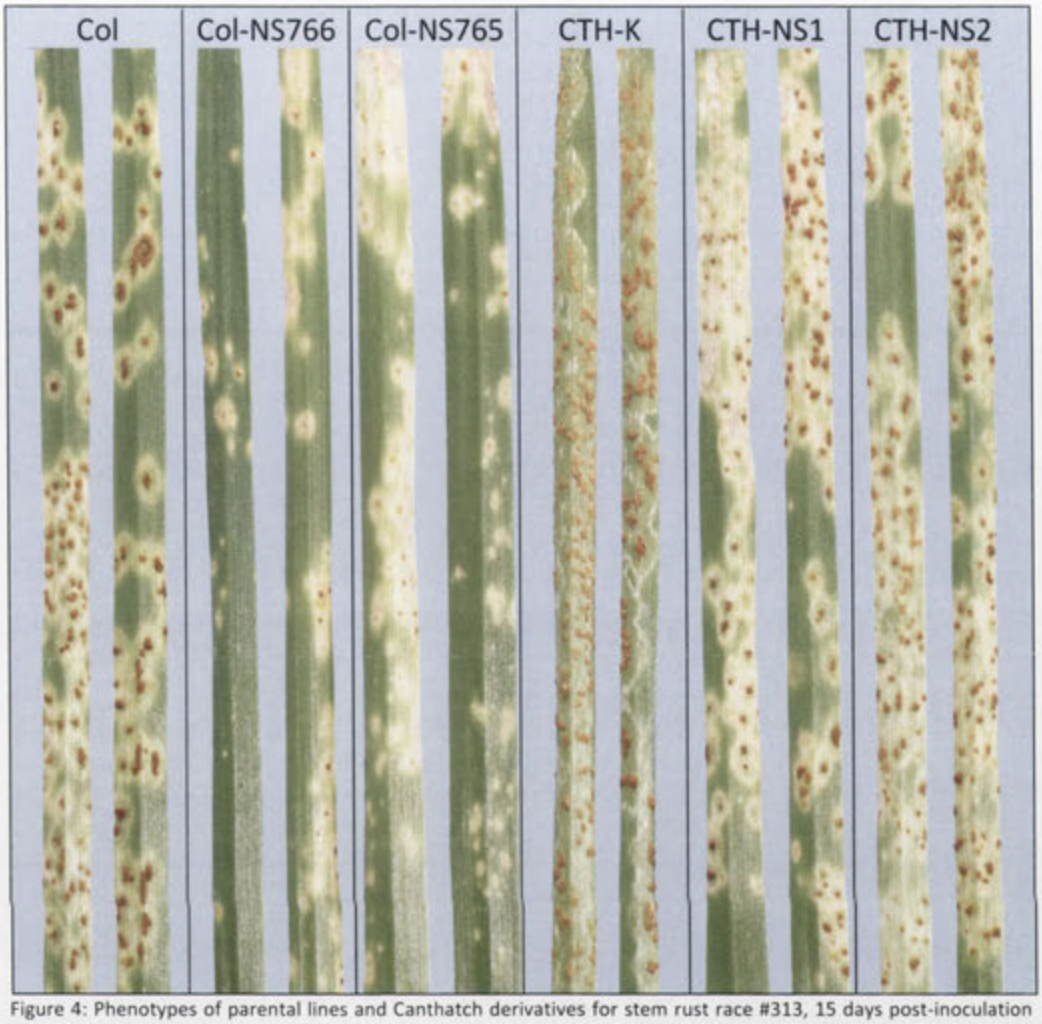


Figure 4: Phenotypes of parental lines and Canthatch derivatives for stem rust race #313, 15 days post-inoculation on seedlings.

I.1.1. Columbus lines

In many cases, the stem rust phenotypes of Col and Col-NS766 were clearly different (Figure 4). Col was moderately susceptible (IT: 2^{+}3), with medium sized uredia surrounded by chlorosis, whereas Col-NS766, as well as Col-NS765, were resistant (IT: ;1), with mainly hypersensitive flecks and a few small uredia surrounded by necrosis and chlorosis.

However, results were variable between experiments, especially in the case of the moderately susceptible Col. This was mainly due to environmental conditions, as the infection on Col was greatly affected by temperature. Incubation at 17-19°C was required to observe good contrast between susceptibility and resistance. At higher temperatures (20-22°C), usually

suitable for stem rust, Col was more resistant. Although variability was sometimes an issue in accurately scoring the plants, phenotypes were repeatable.

I.1.2. Canthatch lines

Several CTH derivatives were also phenotyped using #313. As expected, CTH-K, the CTH line at the origin of the NS mutants (Kerber, 1991), was susceptible (IT: 3^{+}4), with medium to large sized uredia and some chlorosis, whereas CTH-NS1 and CTH-NS2 were moderately resistant (IT: 12^{-}), with small uredia surrounded by chlorosis (Figure 4). The resistance of CTH-NS was weaker than that of Col-NS, however, CTH-K was also more susceptible than Col and comparable to Morocco (IT: 4), which was fully susceptible (Figure 5). Similar to CTH-K, CTH ditelosomic 7DL (CTH-Dt7DL, lacking 7DS) was susceptible (IT: 3), whereas CTH-Dt7DS and CTH nullisomic N7D (CTH-N7D, lacking 7D) were more resistant (IT: 12 and 2^{+}3, respectively) (Figure 5). These results were in agreement with the presence of a suppressor of stem rust resistance on chromosome 7DL (Kerber, 1991; Kerber and Green, 1980). Interestingly, the resistance in CTH-Dt7DS and, especially, in CTH-N7D was not as good as in CTH-NS.

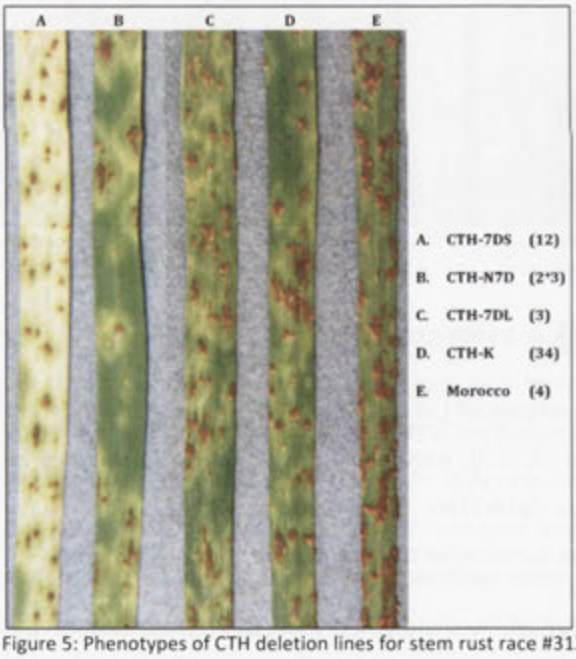


Figure 5: Phenotypes of CTH deletion lines for stem rust race #313. IT score is given in brackets.

I.1.3. Mapping population

Only the first mapping population group (descended from F_1 plants A and B) was tested with #313. Families were phenotyped at least twice, at different time in the year, and F_2

plant zygosity was inferred from the scoring of up to 15 F_3 plants, per assay. F_2 were considered homozygote resistant, if all F_3 were resistant (;1), homozygote susceptible, if all F_3 were susceptible (2*3), or heterozygote, if both resistance and susceptibility were clearly observed in F_3.

Intermediate infection type (IT: ;12) was also observed in individual plants, where hypersensitive flecks and small to medium sized uredia, surrounded by chlorosis, could be seen on the same leaf in variable density. Col and Col-NS766 also sometimes displayed this intermediate phenotype. For this reason, intermediate phenotypes were often ignored and emphasis was given to clear resistant and susceptible phenotypes in the assessment of F_2 zygosity. For some F_3 families, the susceptible phenotype also varied a lot in the same rust infection experiment. Some F_3 displayed the Col-like phenotype (medium sized uredia surrounded by chlorosis) and others, many small uredia surrounded by chlorosis. Figure 6 illustrates this phenomenon for a susceptible F_3 family for which resistant plants were never observed.

Among 128 $F_{2:3}$ families tested, 114 were confidently assigned a phenotypic class. The confidence was based on the reproducibility of the results. However, due to the variability between experiments, some families could have been wrongly scored as segregating, notably the homozygote susceptibles. With the intention to improve accuracy, but also to ensure that the locus conferring resistance to #313 was the same as the one conferring resistance to Ug99, the latter was also used on the mapping population for phenotyping.

Figure 6: Susceptible phenotype variability. Picture of the first leaf originating from 12 individual seedlings from a single susceptible F_3 family at 15 days post-inoculation with stem rust race #313.

I.2. Ug99

According to preliminary tests, Ug99 discriminated Col (IT: 3+4) from Col-NS766 (IT: 0;) more clearly than #313. However, its use is highly restricted and experiments could only be performed at the USDA-ARS, in Minnesota, during wintertime. An assay was also conducted at the Njoro research station of the Kenya Agricultural Research Institute (KARI), Kenya, but failed, as most plants were scored as immune, probably due to a high temperature spike in the glasshouse. Once again, the resistance of plants with Col and CTH backgrounds, seemed greatly affected by environmental conditions. Phenotyping in Minnesota was conducted twice by collaborators Yue Jin and Matt Rouse, for two years (2012 and 2013).

For both years, phenotypes were assessed independently on 10 to 20 plants. The difference between resistant and susceptible infection types was clearer than with #313 and greatly improved the $F_{2:3}$ population phenotyping. However, even with Ug99, phenotypes varied from year to year and from plant to plant in a same experiment, possibly due to uneven inoculations. Interestingly, in 2013 most plants were more resistant than in 2012, reflecting, once again, the impact of environmental conditions on the disease. Despite this, phenotypes were relatively easy to assess.

I.2.1. Columbus and Canthatch lines

In 2012 and 2013, Col was scored as susceptible (IT: 3 and ;3^+ respectively) and Col-NS766 as resistant (IT: ;1 and 0; respectively). The Col-NS765 phenotype was identical to Col-NS766. CTH-K was susceptible in 2012 (IT: 3), but in 2013, many plants were found resistant or, most likely, escaped (IT: 0;). CTH-NS1 and CTH-NS2, only assayed in 2013, were scored as resistant (IT: 0;), although some plants were more susceptible (;13), suggesting that the resistance in the CTH background is less effective than in that of Col. Concerning the CTH deletions lines, for both years CTH-Dt7DL was susceptible (IT: 3^+ and 3^+;/33^+, respectively), whereas CTH-N7D and CTH-Dt7DS were more resistant (IT: from 0; to ;3^+). These results resembled those from #313 and, once again, suggested the presence of a suppressor of stem rust resistance on 7DL.

I.2.2. Mapping population

In 2012, the entire $F_{2:3}$ population was phenotyped, but only the first group (descended from F_1 A and B) and a quarter of the plants descended from F_1 C, could accurately be scored. Overall, the remaining plants were more susceptible, especially those descended from F_1 D and E. These last were assayed at a later date and during a temperature spike that may have prevented some resistance from being expressed. Phenotyping in 2013 was only conducted on the F_3 descended from F_1 A, B and C, due to lack of space.

Among the $F_{2:3}$ families scored in both years, 77% had identical or similar infection types, which was sufficient for mapping. On the other hand, many phenotypes were different between the two years, emphasizing the impact of environmental conditions on the resistance. Most of the variability concerned families that were scored as susceptible or

segregating in 2012 (susceptible year) and segregating or resistant in 2013 (resistant year) respectively. An extreme change from resistant to susceptible was never observed.

I.3. Synthesis

Among the $F_{2:3}$ families confidently scored with #313 and Ug99, 90% had identical or similar phenotypes in response to both races. This suggested that the resistance against both races was the same. #313 could therefore be considered a good substitute for Ug99. Interestingly, all families with different phenotypes were scored as segregating with #313 and susceptible with Ug99. Families were probably scored as segregating with #313, instead of susceptible, because of environmental variations.

Final phenotypes of the $F_{2:3}$ families were based on all experiments with both races. Only those that were consistent were retained for mapping, corresponding to a total of 158 families. 33 were resistant (21%), 86 were intermediate (54%) and 39 were susceptible (25%). These observations fit a one-gene segregation model (χ^2 1:2:1=1.696, df 2, p>0.4).

II. Genotyping

In the attempt to map the 7DL-Sup, molecular markers, known to be on 7DL, were selected and first tested on Col and Col-NS766. Polymorphic markers were then tested on the mapping population.

II.1. Published SSR markers on 7DL

SSRs, that were reported to map to chromosome 7DL (Somers et al., 2004), were the first markers used for mapping the 7DL-Sup. 27 SSRs in total were tested on Col and Col-NS766 (Table Sup. 1) and 4 were found to be polymorphic. 2 markers, co-segregating with each other (wmc0273 and cfa2040), were linked to the rust phenotype, with 10 recombinants out of 118 families. This represented an approximate genetic distance of 8.6 cM.

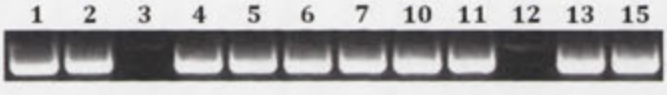
II.2. Published SNP markers on 7DL

Using a recent published chromosome-specific SNP map (Akhunov et al., 2010), 23 pairs of primers, specific to 7DL, were selected for PCR amplification (Table Sup. 2). Their ability to only amplify regions on 7DL was checked using the North Dakota Stem rust Resistant

(NDSR) lines, the NS mutants developed by Williams et al. (1992). One of the mutants, NDSR12, is known to lack the chromosome 7DL, thus, no amplification should occur for this particular mutant, with 7DL-specific primers (Figure 7). All primers were found to specifically amplify 7DL regions, with the exception of one. In addition to their use in detecting primers that are specific to 7DL, the NDSR lines were included in this study in the hope to find polymorphisms between them. Because these lines are independent CTH mutants, finding polymorphisms could facilitate the investigation of the 7DL-Sup.

Interestingly, most of the primers did not amplify any regions of the NDSR3 genome, suggesting that this mutant was also deleted for 7DL. Unlike NDSR12, for which amplification was never seen with 7DL-specific primers, NDSR3 regions were amplified by some primers, indicating a partial deletion. This partial deletion was supported by the SNPs physical order (Akhunov et al., 2010), on which the breaking point of 7DL, in NDSR3, could be visualized (Figure 7B).

A



B

SNP	Rec. Block	NDSR 1	2	3	4	5	6	7	10	11	12	13	15
CD454041	36	+	+	+	+	+	+	+	+	+	-	+	+
BE498849	37	+	+	+	+	+	+	+	+	+	-	+	+
BF474459	38			+							-		
BF201560	40			+							-		
BE490149	41			+							-		
BE494028	43	+	+	-	+	+	+	+	+	+	-	+	+
BE498323	46	+	+	-	+	+	+	+	+	+	-	+	+
BE406943	47	+	+	-	+	+	+	+	+	+	-	+	+
BF482529	48	+	+	-	+	+	+	+	+	+	-	+	+
BE442982	49	+	+	-	+	+	+	+	+	+	-	+	+
BE590469	50	+	+	-	+	+	+	+	+	+	-	+	+
BF485253	53	+	+	-	+	+	+	+	+	+	-	+	+
BE497845	58	+	+	-	+	+	+	+	+	+	-	+	+
BF482403	62	+	+	-	+	+	+	+	+	+	-	+	+
BE444125	67	+	+	-	+	+	+	+	+	+	-	+	+
BG274853	72	+	+	-	+	+	+	+	+	+	-	+	+
BE488496	73	+	+	-	+	+	+	+	+	+	-	+	+
BM137749	77	+	+	-	+	+	+	+	+	+	-	+	+

Figure 7: 7DL specificity of SNP markers. A) Agarose gel of the amplification for BE497845 on NDSR mutants, indicated by numbers. Unlike the other mutants, no amplification was seen on NDSR3 and NDSR12, both lacking 7DL. B) Amplification on NDSR mutants for several SNP markers in the order of the map (Akhunov et al., 2010). Loci are represented by green "+" if present or red "-" if absent. White cases show untested mutants.

16 7DL-specific primers, found to be distal from the NDSR3 breaking point on 7DL, were selected for sequencing in Col and Col-NS766. Only 1 locus, BE497845, was found to be

polymorphic (SNP) between the two parental lines. This SNP, which disrupted a restriction site on Col-NS766 (BsmI), was used as a Cleaved Amplified Polymorphic Sequence (CAPS) marker and tested on the mapping population. It was found to be unlinked to the phenotype.

DISCUSSION

Phenotyping results were consistent between stem rust races #313 and Ug99, indicating that the resistance against both races is common. The resistance in lines with the Col background (Col, Col-NS765 and Col-NS766) was usually better than that in those with the CTH background (CTH-K, CTH-NS1 and CTH-NS2), suggesting than Col has resistance genes and/or components that positively influence the resistance, but are not sufficient for full resistance. Interestingly, the resistance in the CTH deletions lines, CTH-N7D and CTH-Dt7DS, was not as good as observed for the CTH-NS mutants. This could be due to removal of resistance genes present on 7D, or other components affecting plant fitness.

For both races, the infection was greatly affected by environmental conditions. For the sake of accuracy, emphasis was given to clear resistant and susceptible phenotypes, in the assessment of F_2 zygosity. Intermediate phenotypes were often seen in F_3, but as they were also observed in the parental lines, notably Col, most were considered to be due to the environmental effect. However, for some F_3 families, many plants were repeatably scored as intermediate, suggesting a distinct phenotypic class of genetic nature, presumably heterozygosity. Observations using the North American *Pgt* races, C25 and C57, on F_3 descended from the cross between CTH-K and CTH-NS1, showed that the reaction of presumed heterozygote (IT: ;33+), usually approached that of the susceptible group (IT: 2$^+$3$^+$,3$^+$) (Kerber, 1991). The author concluded that the NS allele from CTH-NS1 was recessive, or nearly so. In the present study, intermediate phenotypes were not limited to one infection type, but ranged from resistant to susceptible. This result may indicate that alleles were co-dominant, although with little certainty, due to environmental effect.

Apart from the intermediate phenotypes, the variability of the susceptible phenotype of F_3 families, clearly homozygous susceptible, was troubling (Figure 6). Uneven inoculations could explain this result, since the higher the number of spores on a same leaf, the higher the competition between these individual fungi, resulting in smaller uredia. However, although phenotypes on Col were also variable, such disparity was not observed. This variability could indicate that the mapping population was segregating for more than one gene affecting the

phenotype. These genes could be resistance genes suppressed by the 7DL-Sup in CTH, but not present in Col. On the other hand, the inferred segregation ratio in F_2 was not statistically different from a one-gene segregation (1:2:1), which was expected, as only the 7DL-Sup/NS locus should segregate in this population.

Very few markers on 7DL, retrieved from the literature, were found to be polymorphic between Col and Col-NS766. However, since these markers were based on different genotypes, it was likely that most of them were inappropriate for this study. In addition, after five backcrosses, polymorphism was expected to be low, all the more since Col and CTH-NS, which was at the origin of Col-NS766, are closely related. The aim of this study was simply to coarsely map the 7DL-Sup, thus published markers were expected to be sufficient. Nevertheless, two co-segregating SSR markers were still found to be linked to the phenotype. However, the genetic distance of 8.6 cM was too far for physical mapping, but provided an idea of the location of the locus being mapped. According to the SSR wheat consensus map (Somers et al., 2004), these markers are located on the distal part of chromosome 7DL, which implies that the locus is also on the distal part. Subsequent experiments that are discussed in the next chapter showed that this locus did not correspond to the 7DL-Sup locus, but a resistance locus on 7AL.

Although not polymorphic between Col and Col-NS766, the SNP markers indicated the rough location of the 7DL-Sup. Indeed, one of the NDSR mutants, NDSR3, was found to be partially deleted for 7DL and the breaking point was approximately mapped. This deletion was probably at the origin of the resistance in NDSR3 (Williams et al., 1992), hence the 7DL-Sup may be located below the breaking point.

As the published markers were not sufficient to clearly map the 7DL-Sup, other approaches were investigated. The use of new technologies for marker discovery and mapping is discussed in the next chapter.

MATERIALS AND METHODS

Plant materials

All Columbus and Canthatch lines/mutants were provided by E. R. Kerber, Agricultural Canada Research Station in Winnipeg. The $F_{2:3}$ mapping population was developed from 5 F_1 of the cross Col x Col-NS766. The NDSR mutants were provided by S.S Xu, USDA-ARS, Fargo,

North Dakota. The isolation of DNA from leaves was carried out as detailed in Lagudah et al. (1991).

Evaluation of seedling response to stem rust

Resistance to wheat stem rust *Puccinia graminis* f. sp. *tritici* was evaluated with two races. For #313 (pathotype 34-1,2,3,5,6,7), plants were scored at the Plant Breeding Institute, University of Sydney. Seedlings were grown in growth chamber (20-22°C, natural light cycle - max 170 µmol/m2/s) until 2-leaf stage (≈2 weeks). Inoculation was done by spraying the plants with spores suspended in oil. Plants were let to dry for 30 min and incubated for 2 days in incubation cabinets (20-22°C, high humidity, natural light cycle - max 85 µmol/m2/s). Plants were moved back to growth chamber at lower temperature (17-19°C) until ready for scoring (≈12 days). For Ug99 (TTKSK), plants were scored at the USDA-ARS Cereal Disease Laboratory in St. Paul, Minnesota, following the method detailed in Rouse et al. (2012). Chi-square goodness-of-fit tests were used to evaluate the observed resistance segregations with the expected ratios.

Marker analyses

SSR and SNP primer sequences were retrieved on GrainGenes 2.0 (http://wheat.pw.usda.gov/GG2/index.shtml). SSR assays were done using the method of (Rampling et al., 2001) on 3130xL Genetic Analyzer (Applied Biosystems®) and analysed with the SoftGenetics GeneMarker® software version 1.97. SNP assays were evaluated by PCR with standard conditions and visualized by electrophoresis on agarose gel. Linkage analysis was performed using the Kosambi map function (Kosambi, 1943).

REFERENCES

Akhunov, E.D., Akhunova, A.R., Anderson, O.D., Anderson, J.A., Blake, N., Clegg, M.T., Coleman-Derr, D., Conley, E.J., Crossman, C.C., Deal, K.R., et al. (2010). Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes. BMC Genomics *11*, 702.

Feuillet, C., Travella, S., Stein, N., Albar, L., Nublat, A., and Keller, B. (2003). Map-based isolation of the leaf rust disease resistance gene *Lr10* from the hexaploid wheat (*Triticum aestivum* L.) genome. Proc. Natl. Acad. Sci. *100*, 15253–15258.

Kerber, E.R. (1991). Stem-rust resistance in "Canthatch" hexaploid wheat induced by a nonsuppressor mutation on chromosome 7DL. Genome *34*, 935–939.

Kerber, E.R., and Green, G.J. (1980). Suppression of stem rust resistance in the hexaploid wheat cv. Canthatch by chromosome 7DL. Can. J. Bot. *58*, 1347–1350.

Kosambi, D.D. (1943). The estimation of map distances from recombination values. Ann. Eugen. *12*, 172–175.

Krattinger, S., Wicker, T., and Keller, B. (2009). Map-based cloning of genes in Triticeae (wheat and barley). In Genetics and Genomics of the Triticeae, G.J. Muehlbauer, and C. Feuillet, eds. (New York, NY: Springer US), pp. 337–357.

Lagudah, E.S., Appels, R., and McNeil, D. (1991). The Nor-D3 locus of *Triticum tauschii*: natural variation and genetic linkage to markers in chromosome 5. Genome *34*, 387–395.

McIntosh, R.A., Wellings, C.R., and Park, R.F. (1995). Wheat rusts an atlas of resistance genes (East Melbourne, Australia; Dordrecht; Boston: CSIRO ; Kluwer Academic Publishers).

Rampling, L.R., Harker, N., Shariflou, M.R., and Morell, M.K. (2001). Detection and analysis systems for microsatellite markers in wheat. Aust. J. Agric. Res. *52*, 1131–1141.

Rouse, M.N., Nava, I.C., Chao, S., Anderson, J.A., and Jin, Y. (2012). Identification of markers linked to the race Ug99 effective stem rust resistance gene *Sr28* in wheat (*Triticum aestivum* L.). Theor. Appl. Genet.

Semagn, K., Bjørnstad, Å., and Ndjiondjop, M.N. (2006). An overview of molecular marker methods for plants. Afr. J. Biotechnol. *5*.

Somers, D.J., Isaac, P., and Edwards, K. (2004). A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). Theor. Appl. Genet. *109*, 1105–1114.

Stein, N., Feuillet, C., Wicker, T., Schlagenhauf, E., and Keller, B. (2000). Subgenome chromosome walking in wheat: a 450-kb physical contig in *Triticum monococcum* L. spans the *Lr10* resistance locus in hexaploid wheat (*Triticum aestivum* L.). Proc. Natl. Acad. Sci. *97*, 13436–13441.

Wicker, T., Stein, N., Albar, L., Feuillet, C., Schlagenhauf, E., and Keller, B. (2001). Analysis of a contiguous 211 kb sequence in diploid wheat (*Triticum monococcum* L.) reveals multiple mechanisms of genome evolution. Plant J. *26*, 307–316.

Williams, N.D., Miller, J.D., and Klindworth, D.L. (1992). Induced mutations of a genetic suppressor of resistance to wheat stem rust. Crop Sci. *32*, 612–616.

Zhang, Y.-M., and Xu, S. (2004). Mapping quantitative trait loci in F2 incorporating phenotypes of F3 progeny. Genetics *166*, 1981–1993.

SUPPLEMENTAL

SSR	Forward Sequence 5'-3'	Reverse Sequence 5'-3'
barc0053	GCGTCGTTCCTTTGCTTGTACCAGTA	GCGCGTCCTTCCAATGCAGAGTAGA
barc0111	GCGGTCACCAGTAGTTCAACA	GCGTATCCCATTGCTCTTCTTCACTAAC
barc0121	ACTGATCAGCAATGTCAACTGAA	CCGGTGTCTTTCCTAACGCTATG
barc0172	GCGAAATGTGATGGGGTTTATCTA	GCGATTTGATTTAACTTTAGCAGTGAG
barc0026	GCGCTGGGTAAAAAGTGAAATTC	TGCAAGTGGAGGGGGAGGCGAGAG
cfa2040	TCAAATGATTTCAGGTAACCACTA	TTCCTGATCCCACCAAACAT
cfd0025	CATCGCTCATGCTAAGGTCA	CGTGTCTGTTAGCTGGGTGG
cfd0069	AAATACCTTGAATTGTGAGCTGC	TCTGTTCATCCCCAAAGTCC
gdm0046	TGTGTTGGCCTTGTGGTG	CTACCCAATGCATCCCCTTA
gdm0067	AAGCAAGGCACGTAAAGAGC	CTCGAAGCGAACACAAAACA
gwm0037	ACTTCATTGTTGATCTTGCATG	CGACGAATTCCCAGCTAAAC
gwm0428	CGAGGCAGCGAGGATTT	TTCTCCACTAGCCCCGC
gwm0437	GATCAAGACTTTTGTATCTCTC	GATGTCCAACAGTTAGCTTA
MAG892	TCTTCTTCCTCTCTCAAGGCA	CGTAACCAAGCTGGGAAGAG
PSP3003	GATCGACAAGGCTCTAATGC	CAGGAGGAGAGCCTCTTGG
wmc0014	ACCCGTCACCGGTTTATGGATG	TCCACTTCAAGATGGAGGGCAG
wmc0157	CTTGATCCAAGTGGTTCTTTCC	TCCAAATGTTTGCGAAACCTGA
wmc0166	ATAAAGCTGTCTCTTTAGTTCG	GTTTTAACACATATGCATACCT
wmc0221	ACGATAATGCAGCGGGGAAT	GCTGGGATCAAGGGATCAAT
wmc0273	AGTTATGTATTCTCTCGAGCCTG	GGTAACCACTAGAGTATGTCCTT
wmc0473	TCTGTTGCGCGAAACAGAATAG	CCCATTGGACAACACTTTCACC
wmc0671	GTACGTCAAAGAAAGAGAATTACCTC	CTCAGAGATATATCTTCGTTGTCAGT
wmc0797	CGAAACCCTAGATGAAGC	ACACAACCACAGGTGAGTTGTTCT
wmc0824	CCGATGAACTTAAAAGTACCACCTG	CATGGATTGACACGATTGGC
wmc0094	TTCTAAAATGTTTGAAACGCTC	GCATTTCGATATGTTGAAGTAA
wmc0438	GACCGTTGGGCTGTATAGCATT	CTCTGACAGTGGTGGAGCTTGA
wmc0488	AAAGCACAACCAGTTATGCCAC	GAACCATAGTCACATATCACGAGG

Table Sup. 1: List of SSR markers tested on the mapping population (Somers et al., 2004).

SNP	Fwd Name	Fwd Sequence 5'-3'	Rev Name	Rev Sequence 5'-3'
BE398591	BE398591D_F1	ACAGTGCGTGATATCATGTTG	BE398591_cpR1	CTCCTTGCTGAAAAAGGTGG
CD454041	CD454041D_F2	ATTCTTCAAGCATGATAACAAAG	CD454041_cpR1	TGCTTTGATGCAGCACTAGG
BE498849	BE498849D_F3	CTTGTCTCAGATTTGTCTAGATAT	BE498849_cpR1	CATTGGCACCTCCTGTACCT
BF474459	BF474459D_F1	CCTTGAATGATGGCTGACATTA	BF474459_cpR1	CCTGCTGAGCCTTGAACATT
BE444691	BE444691D_F2	GTGATGATGATCTTTGCTGTAG	BE444691_cpR1	TTTGATATCCCACCCGACAT
BF201560	BF201560D_F3	TGAAGGTATAGATGACAAGACTA	BF201560_cpR1	CAGCTGCTTCCAAGAGCATC
BE490149	BE490149_cpF1	AGCCAAGGAGCAAGTTGAAA	BE490149D_R3	CTTGACAGGTAGGCTTCTAC
BE494028	BE494028D_F1	GAAGGCATGAAGAGAGGAACG	BE494028_cpR1	GGTTTCATCACTGAACCGCT
BE443312	BE443312_cpF1	AGCAAAGCCCTATACGCTGA	BE443312D_R1	AACATACATAGACCTCAGTAACA
CD373508	CD373508D_F2	CATTAAGAAATGCCAAAACTTATAG	CD373508_cpR1	GAGCTTATCCAGAGGGGTCC
BE498323	BE498323D_F1	TTGCTTTCTATTTTATTTTGTTGG	BE498323_cpR1	CAGATCCATAATCTCCCCCA
BE406943	BE406943_cpF1	TGAGACTCCTGATGGTGCTG	BE406943D_R1	ATGCTTCAAGGGCCCTCTTC
BF482529	BF482529D_F1	GAAAATACAAGGGTGAACATGGA	BF482529_cpR1	CGTGAACTTCTTCACGTATCCA
BE442982	BE442982_cpF1	TGGGGAATGAAGTCGGTAAG	BE442982D_R1	GAAAGAAAAACTTGGTCTTA
BE590469	BE590469D_F2	TAAAATTTGGAAAAGAAGCTCAAA	BE590469_cpR1	ACTTGACAACTTTCGGACGG
BF485253	BF485253D_F2	ATTCTTGATTTCACTTGTTACTAT	BF485253_cpR1	CAATGTATGAGGCGTGTTCG
BE591273	BE591273_cpF1	CAACCTTTTGTGCTGGGAGT	BE591273D_R2	TGCTTGTCTGAATCAGCAATG
BE497845	BE497845D_F1	CATCTTCTCGAGTTTGCACC	BE497845_cpR1	CTTTGCCTTTCGAGGAACTG
BF482403	BF482403_cpF1	CTTACTGTGGATCTTGGCCC	BF482403D_R2	GTGACTCACCAATTGTACCC
BE444125	BE444125_cpF1	AAACCGTTCCTGTTTTGCAC	BE444125D_R2	TAACCGATACAACGAAACAAG
BG274853	BG274853D_F3	CTGGAATTTAATTTGCATGGTC	BG274853_cpR1	GTTGCTACCGTGCACTTTGA
BE488496	BE488496_cpF1	GGTGGATGTTCTGCTGGATT	BE488496D_R1	TCCATCCACATTCAGCTCG
BM137749	BM137749_cpF1	GAGGGCATAAACCGACTGAA	BM137749D_R1	CCTCTAGCAATGCAAAGTTC

Table Sup. 2: List of SNP markers tested on the mapping population (Akhunov et al., 2010).

Chapter 3

High-throughput genotyping locates a locus on chromosome 7AL linked to the stem rust resistance in Columbus-NS766

ABSTRACT

In order to obtain new markers closely linked to the 7DL suppressor, high-throughput genotyping techniques were used on Columbus and Columbus-NS766, as well as on bulked segregants. Thousands of new polymorphisms were found, with several hundred being putatively linked to the stem rust resistance. Mapping of the segregating population unexpectedly revealed close linkage between the polymorphic markers and a resistance locus on chromosome 7AL. No significant association was detected between the markers and chromosome 7DL, presumed to carry the non-suppressor locus in Columbus-NS766. It appears that a stem rust resistance locus on 7AL was introgressed from Canthatch-NS in the course of developing the Columbus backcross derivative, Columbus-NS766. Many markers were found to be part of a cluster that showed linkage with the 7AL resistance locus.

List of major abbreviations

7DL-Sup Canthatch 7DL suppressor; **9k chip** 9k SNP chip; **BR** resistant bulk; **BS** susceptible bulk; **BSA** bulked segregant analysis; **Col** Columbus; **Col-NS** Columbus-NonSuppressor; **CS** Chinese Spring; **CTH** Canthatch; **CTH-NS** Canthatch-NonSuppressor; **CTH-Dt7DL** CTH ditelosomic 7DL; **CTH-Dt7DS** CTH ditelosomic 7DS; **CTH-N7D** CTH nullisomic 7D; **GBS** genotyping by sequencing; **GSS** genome survey sequencing; **KASP** KBiosciences Competitive Allele - Specific PCR genotyping system; **NDSR** North Dakota Stem rust Resistant; **NS** non-suppressor; **PE** paired-end; **Pgt** *Puccinia graminis* f. sp. *tritici*; **RAD-Seq** restriction-site associated DNA sequencing; **SNPs** single nucleotide polymorphisms; **SSR** simple sequence repeats.

INTRODUCTION

Given the paucity of marker, that were shown to be polymorphic between Columbus (Col) and Columbus-NS766 (Col-NS766), of which only two were linked to stem rust resistance, new markers had to be found. Up until recently, developing new markers was a laborious and time-consuming task, often requiring preliminary genomic knowledge. With constant improvement in technologies, this process has become much easier and faster, enabling the discovery of thousands of markers in a short time (Henry et al., 2012). In order to find new markers for genetic mapping of the 7DL suppressor (7DL-Sup) locus, genome wide genotyping using new technologies were undertaken on Col and Col-NS766. Two techniques were exploited: i) genotyping of nearly 9,000 single-nucleotide polymorphisms (SNPs) using

Illumina® iSelect SNP assay and ii) genotyping by sequencing (GBS) on complexity reduced genome. New markers were usually converted into KASP assays and used for mapping the segregating population.

I. Wheat 9k SNP chip

The wheat 9k SNP chip (9k chip) is an Infinium iSelect BeadChip recently developed, which enables genotyping of nearly 9,000 gene-associated SNPs (Cavanagh et al., 2013). These SNPs were discovered in transcriptomes from wheat cultivars representing the major breeding programs of Australia, eastern and western Asia, Europe and North America, and are therefore fairly common and highly heritable in wheat. With the simultaneous genotyping of thousands of SNPs, this chip is ideal for quick genome wide mapping studies.

I.1. Principle

The chip contains 8,632 beads, surface of which is covered by multiple copies of a unique probe. Their sequence directly flanks the SNP without the SNP itself. Genomic DNA is first amplified and fragmented into small fragments, which are then hybridized to the appropriate probe. Following a single-base extension, a nucleotide corresponding to the DNA sequence is added to the probe, and stained with different fluorophores. Finally, the emitted fluorescence is measured and genotypes are determined based on the fluorescence ratio. For each SNP assayed, alleles are attributed on the basis of the two output values:

- **The normalized Theta (θ) value**: represents the allele frequency. A θ value of 0 indicates that only the A allele is detected, whereas 1 indicate that only the B allele is detected. Any values between could indicate heterozygosity (A and B for a same locus) or the presence of multiple loci, possibly homoeologous, with different alleles across the genome (A for one locus and B for another).
- **The normalized R value**: represents the intensity of the signal. R values near 0 indicate that the locus is absent.

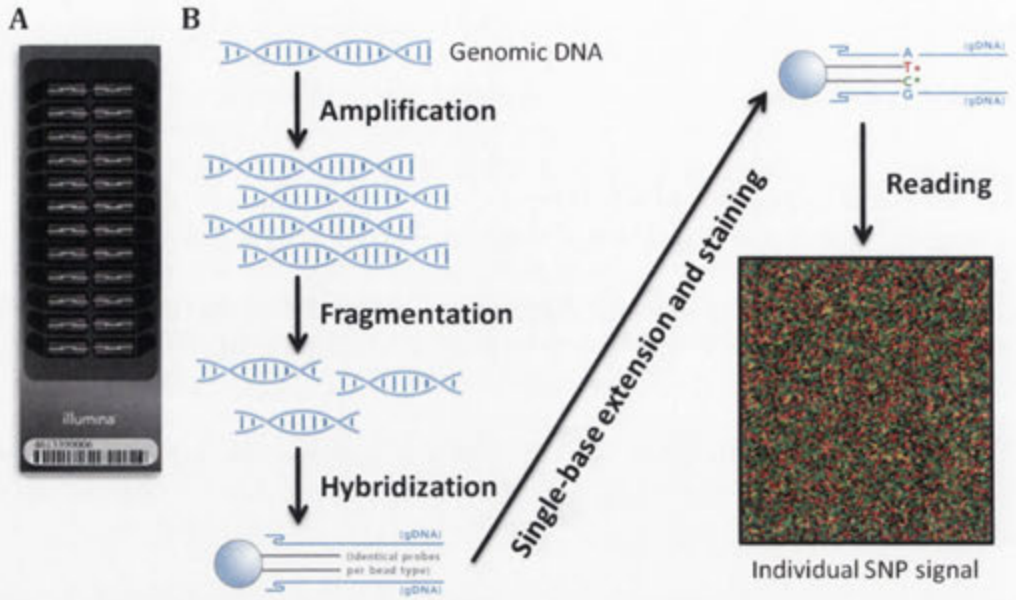


Figure 1: Infinium iSelect. A) Picture of a 24 assays beadchip. B) Genotyping principle. Illustration inspired from the Infinium® II Assay Workflow.

II. Genotyping by sequencing

In a further attempt to develop markers tightly linked to the 7DL-Sup locus, the newly developed GBS methodology was used. Like the 9k chip, GBS allows for simultaneous SNP discovery and genotyping, but SNPs are specific to the sequenced individual and not pre-determined. It is efficient and cost-effective and allows the development of high-density genetic maps (Baird et al., 2008; Poland et al., 2012). In order to quickly find closely linked markers, GBS using the restriction-site associated DNA (RAD) method was done on bulked segregants.

II.1. Principle

II.1.1. Bulked Segregant Analysis

Bulked Segregant Analysis (BSA) is an efficient mapping technique for rapidly identifying genomic regions affecting a trait of interest (Michelmore et al., 1991). It is based on the analysis of DNA pooled from individuals sharing a particular trait and coming from a segregating population. By doing so, all alleles present in the population are mixed, except at the region responsible for the phenotype. Consequently, two bulks differing for a particular trait

are virtually identical except at the region surrounding the gene conferring the trait. This gene can only come from one parental line in a particular bulk (e.g. from the resistant parent in the resistant bulk, and from the susceptible parent in the susceptible bulk). By assessing allele frequencies in both bulks, it is therefore possible to quickly associate a genomic region to the trait of interest (Figure 2).

Because the technique greatly facilitates the identification of any specific genomic region, BSA has been widely used in genetic mapping and successfully lead to the development of closely linked markers in many species including wheat (Asad et al., 2012; Ma et al., 2011). So far, attempts to map the 7DL-Sup were based on relatively random markers, without preliminary knowledge on their linkage to the gene. By using BSA, markers developed through the sequencing of bulks should be linked to the gene.

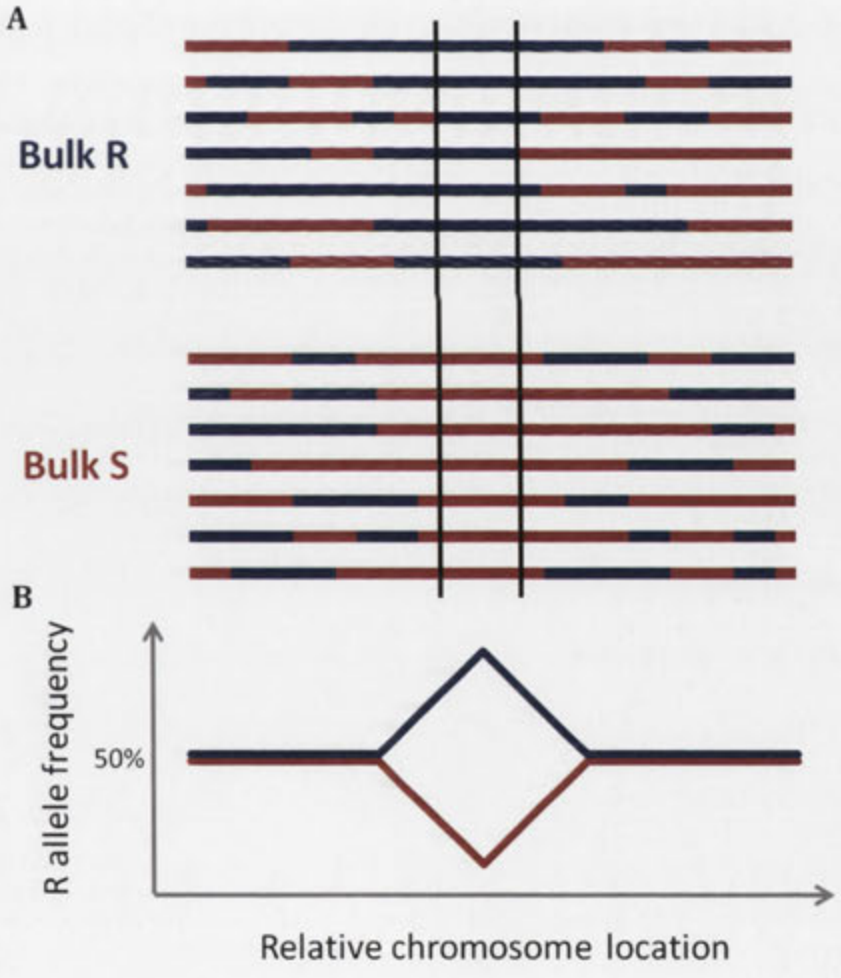


Figure 2: Alleles in bulked segregants. A) Chromosome fragments coming from the resistant (blue) or the susceptible (red) parent in the different bulks. The central region represents the locus conferring the trait of interest. B) Frequency of alleles coming from resistant parent in both bulks. The closer to the locus of interest, the greater the allele frequencies difference between the bulks.

II.1.2. Restriction Site Associated DNA Sequencing

For many species, whole genome sequencing can be current practice, as their genome is small enough to do so. However, this is not practicable yet in wheat due to the genome size and the number of repetitive regions. It is, therefore, important to dramatically reduce the genome complexity before proceeding to sequencing. Several methods are available (Poland and Rife, 2012), such as restriction site associated DNA sequencing (RAD-Seq) (Baird et al., 2008; Davey and Blaxter, 2011). It is based on the sequencing of regions adjacent to restriction sites (Figure 3). DNA is first cleaved by a restriction endonuclease, before being randomly sheared (e.g. by sonication), creating fragments with two distinct ends, or paired-end (PE); PE2 at the restriction site and PE1 at the random shear site. Apart from getting twice as much data as with single-end sequencing, paired-end sequencing allows local *de novo* assemblies, which is especially useful for primer design when a reference genome is not available. The fragments size can be controlled during shearing and by post-selection (e.g. on agarose gel). Both fragment ends are ligated with special adapters; which contain molecular identifiers and various barcodes, allowing sample multiplexing.

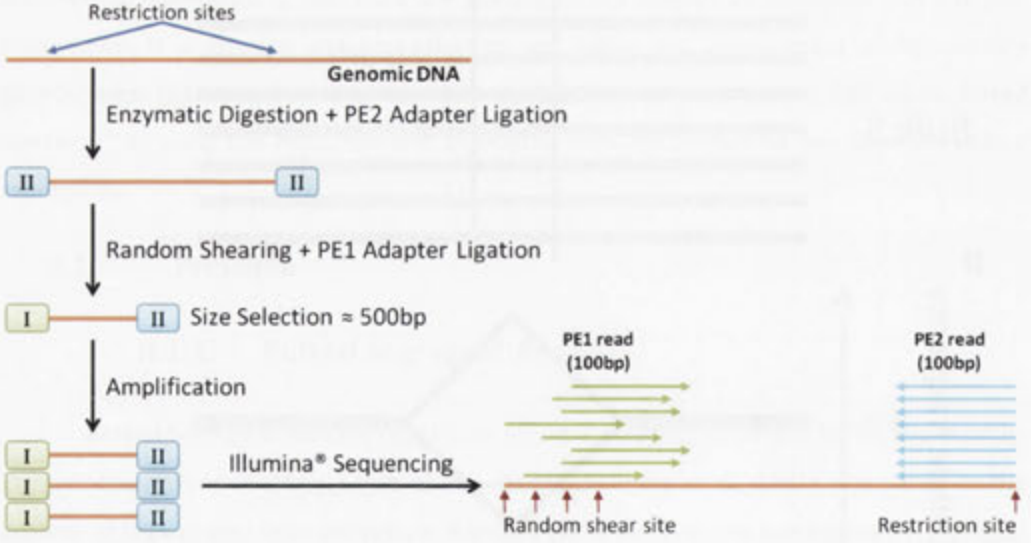


Figure 3: RAD-Seq principle. Genomic DNA is first digested with a restriction enzyme, and then randomly sheared into fragments of controlled size. Both ends are ligated with adapters, sequences of which are used for amplifying the fragments. Upon Illumina sequencing, 100bp reads are generated for each fragment end.

Methylation-sensitive enzyme

The choice of the restriction enzyme can considerably affect the complexity, as the frequency of recognized sites can vary between enzymes. Furthermore, cleavage by some enzymes is inhibited by methylation of the DNA target sequence. Those are particularly interesting, as low-copy and gene rich regions are often hypo-methylated, whereas repetitive elements are hyper-methylated (Bennetzen et al., 1994) (Figure 4). Use of Methylation-sensitive enzymes, such as *AatII* and *PstI*, were shown to be very effective in the enrichment of sequences for gene regions in wheat, reducing repetitive regions by 80% (Fellers, 2008).

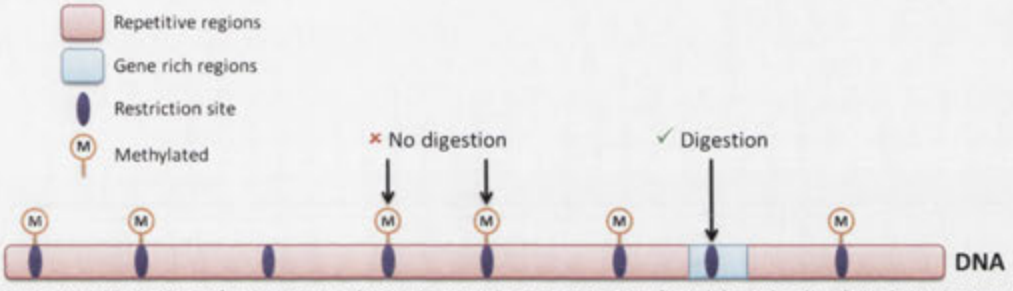


Figure 4: Methylation and genome complexity. A particular restriction site (purple) can be distributed evenly across the genome, which is mainly constituted of repetitive regions (red) in wheat. As these regions are often more methylated than gene rich regions (blue), the use of methylation-sensitive restriction enzymes allows the removal of most repetitive regions.

II.2. Strategy

As seen above, RAD-Seq and BSA are both powerful methods for genetic mapping. The former allows the rapid discovery of molecular markers, and the latter indicates their linkage to the trait of interest. The combination of both methods enables the detection of markers directly in the region of interest (Baird et al., 2008). Similar approaches using RNA-Seq, have also been shown successful for the discovery of new closely linked markers (Liu et al., 2012; Trick et al., 2012). Like all high-throughput technologies, RAD-Seq on BSA includes various challenges, notably the need for a good bioinformatics pipeline.

Illumina® sequencing allows the simultaneous generation of millions of reads. In order to compare them, each read must be first associated with a unique locus, by aligning them against a reference. This reference can consist in the published reference genome, if only this one had already been sequenced, which is still too rare especially for non-model species. Fortunately, this is becoming possible in wheat, thanks to recent advances in the wheat genome survey sequencing (GSS) (see below). Another possibility is to build one by *de novo*

assembling the reads or easier, by "stacking" reads that matches perfectly (Catchen et al., 2011). This last is particularly appropriate for RAD-Seq, as all PE2 reads (from the restriction site end) start at the same position and can thus easily pile up, as long as they come from the same genomic region. In contrast, PE1 reads cannot be stacked as efficiently, as their starting position varies from each other.

Once the alignments done, SNPs are detected between each read and the corresponding aligned region. They are then compared between samples and linkage is evaluated. SNPs are considered putatively linked when alleles are different between bulks and in phase with the appropriate parent (Table 1). The clearest results are for SNPs that are homozygous in the parents, and for which a single allele is detected in each bulk. These SNPs are classified as genome specific polymorphism (GSP), as opposed to multiple sequence variants (MSV), for which multiple alleles were detected in one bulk and the corresponding parent.

SNP	Parents		Bulks		Classification	Confidence
Type	S	R	S	R		
SNP 1	A	C	A	C	GSP	1
SNP 2	A	.	A	C	GSP	2
SNP 3	.	.	A	C	GSP	3
SNP 4	A	A/C	A	A/C	MSV	1

Table 1: Linked SNP classification. A and C represent the two variants of a SNP. Dots represent missing data. Depending on the number of allele for each genotype, SNPs are considered GSP (genome specific polymorphism) or MSV (multiple sequence variants). The more genotypes are characterized, the better the confidence. S and R represent susceptibility and resistance to stem rust, respectively.

The wheat genome survey sequencing

The wheat GSS is undertaken by the International Wheat Genome Sequencing Consortium (IWGSC; http://www.wheatgenome.org), which aims to develop a high quality reference genome for bread wheat. It is based on the sequencing of all 21 sorted chromosomes from the cultivar *Chinese Spring* using Illumina and 454 technologies, as well as BAC sequencing. Although not complete yet, millions of genomic contigs of various sizes from 200 bp to 130 kbp have been generated. Moreover, as it is developed from sorted chromosomes of the cultivar Chinese Spring, the greatest advantage of the GSS is the opportunity to attribute any aligned sequence to a chromosome.

III. KASP genotyping

Until recently, SNP genotyping was mainly based on techniques relying on restriction enzymes to generate cleaved amplified polymorphic sequence (CAPS) which are time consuming, expensive and not appropriate for high-throughput analysis. Allele specific (AS) PCR is a cheaper technique that relies on the specific amplification of a particular allele using oligonucleotides directly designed on SNPs. Although supposed to be specific, one nucleotide difference is often insufficient to distinguish between two alleles with conventional PCR. In addition, homoeologues loci can also complicate the analysis.

The KBiosciences Competitive Allele-Specific PCR genotyping system (KASP) is a very good alternative for SNP genotyping, especially in complex species such as wheat (Allen et al., 2011). Although this technique has been available for some years now, it is still widely unknown. It relies on the competition of AS oligonucleotides, which minimises non-specific amplification (Figure 5). Each AS-primer is coupled with a unique tail sequence which interacts, upon amplification, with a specific fluor-labeled oligo. To ensure specific amplification, both AS-primers are used in the same reaction, which favours the amplification of the correct allele. Finally, the fluorescence signals are measured and alleles are determined (Figure 6). Several advantages can be attributed to this technique: only one PCR required; increased allele specificity; direct detection of heterozygotes and it works with multiple copies present.

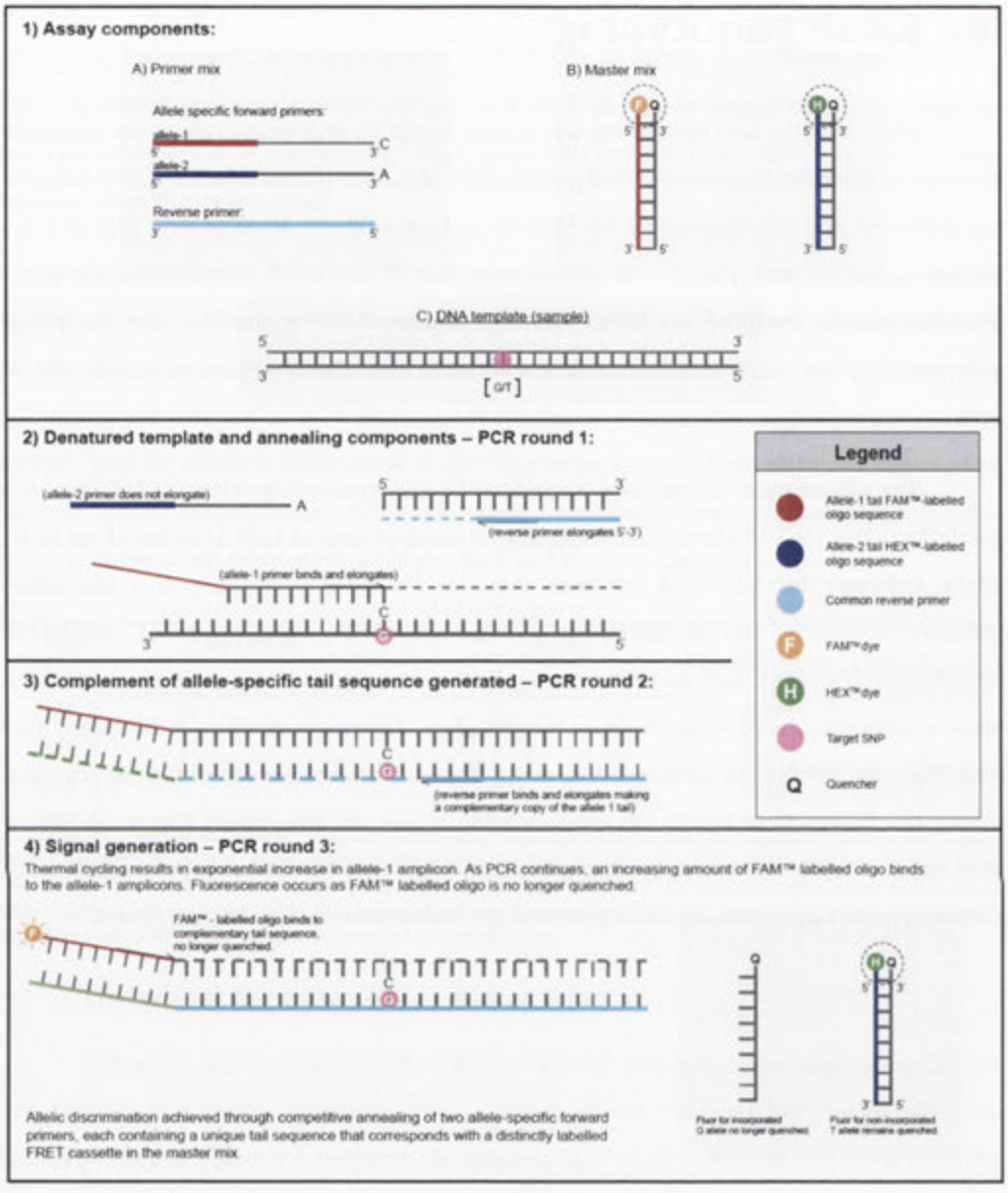


Figure 5: Principle of KASP assay genotyping. Source: http://www.lgcgenomics.com.

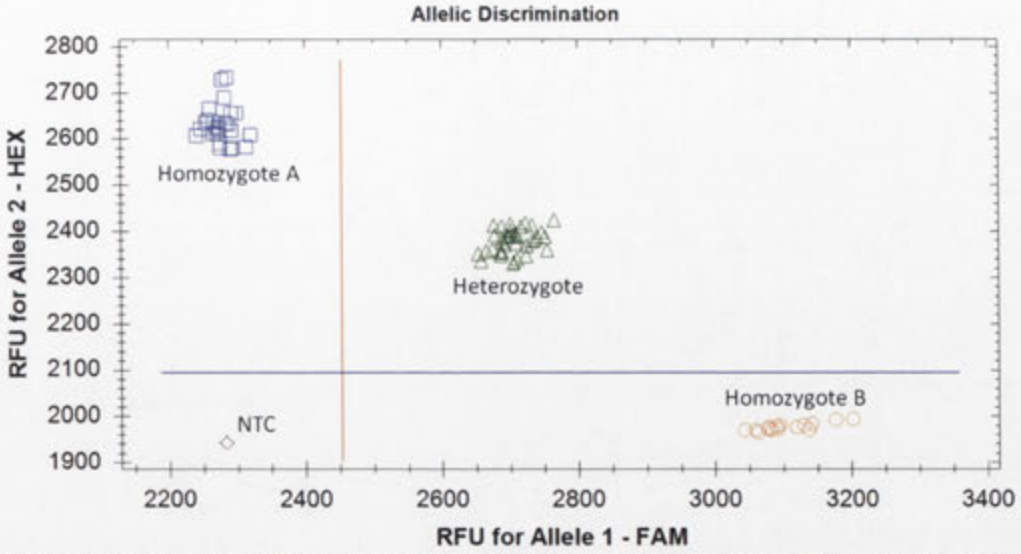


Figure 6: Example of KASP genotyping using Bio-Rad CFX Manager. Amplified alleles are detected through the measure of the relative fluorescence units (RFU) from both FAM and HEX fluorophores, each associated to a particular allele. Only one fluorescent signal is emitted in homozygotes, whereas both fluorescence signals are detected in heterozygotes. If no amplification, no signal is detected like in the non-template control (NTC).

RESULTS

I. Wheat 9k SNP chip

In order to find new polymorphisms between Col and Col-NS766, DNA from both lines were loaded on the chip. DNA from each of the 12 North Dakota Stem rust Resistant (NDSR) lines, which are mutated for the 7DL-Sup locus (Williams et al., 1992), were also loaded, with the intention of detecting SNPs that are located on chromosome 7DL. Two of these mutants carried large deletions for 7DL, therefore any difference between them and the other mutants would suggest that the SNP was located on 7DL. Moreover, the NDSR lines were included in the hope to find polymorphism between them, which could facilitate the investigation of the 7DL-Sup. All lines were genotyped for the 8,632 SNP assays. Depending on the results between Col and Col-NS766, as well as between the NDSR mutants, various classifications were assigned to each locus (Figure 7).

A threshold for the signal intensity R was first evaluated using the variation between the four replicates of Col and Col-NS766. With R close to 0, the θ values between replicates were very dissimilar, as reflected by the standard deviation (Figure Sup. 1C and D). Loci were considered absent if R was below 0.14 (Figure 7B).

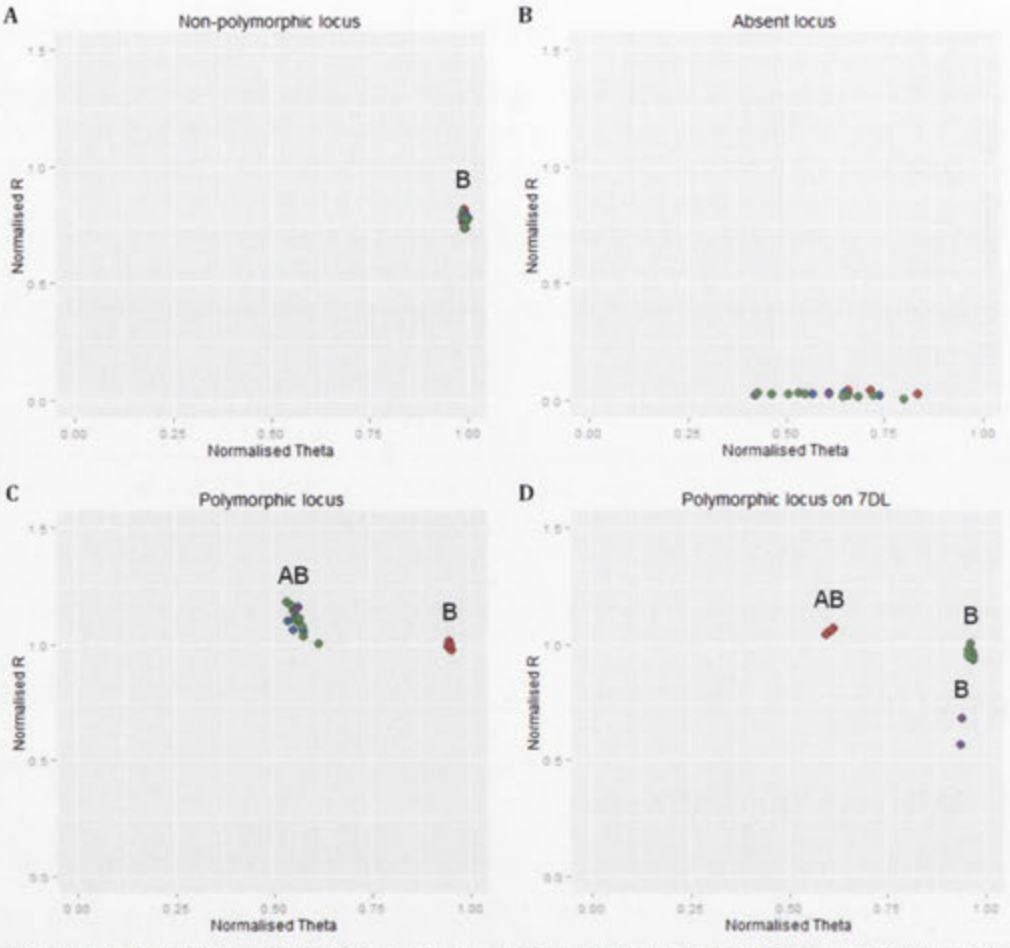


Figure 7: Examples of locus classification. 9k chip output for 4 SNP assays with different classifications. Dots represent Col (red), Col-NS766 (blue), NDSR+ (NDSR lines with 7DL, green) and NDSR- (NDSR lines without 7DL, purple). The letters above each group of samples represent the detected alleles. Note: samples with both alleles detected (AB) could be heterozygotes or possess multiple loci with different alleles. Multiple loci with the same allele are illustrated in the example D, where the B allele was detected on both NDSR+ and NDSR-, but the signal intensity was weaker on the latter, as if one copy was missing (presumably on 7DL).

I.1. Detection of polymorphism between Col and Col-NS766

Polymorphism between Col and Col-NS766 was assessed by comparing the θ values (polymorphic alleles) or R values (missing alleles) (Figure 7C and D). Out of the 8,632 SNPs assayed, 247 loci were found polymorphic and 87% of them were due to an allelic variation (change in θ). Few polymorphic loci were found with a difference of θ values approaching 1 (Figure 8), signalling the presence of multiple copies for most loci. This reflected the chip sensitivity for clearly distinguishing polymorphism, even with multiple copies.

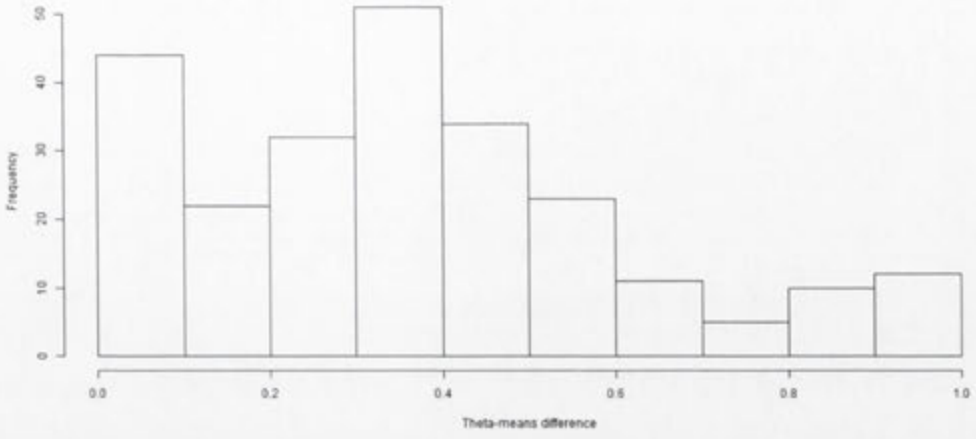


Figure 8: Histogram showing the distribution of the differences between the means of θ values of Col and Col-NS766 for the polymorphic loci. With a difference approaching one, Col and Col-NS766 are considered homozygous for each locus.

I.2. Detection of loci on 7DL

Because all NDSR mutants share the same genetic background, they were expected to give very similar results. This was verified by testing the Pearson's correlation of the θ and R values between NDSR mutants. For both values, all mutants were well correlated between each other (minimum r^2= 0.987 and 0.966 for θ and R, respectively), indicating that most loci were identical in all NDSR mutants. Nevertheless, NDSR3 and NDSR12, which lack most of the 7DL chromosome arm, were expected to be slightly different from the other mutants. In order to make sure of this, a hierarchical clustering analysis of all NDSR mutants was performed, using the θ or R values (Figure 9). As expected, NDSR12 was dissimilar from the main group of NDSR mutants. Surprisingly, NDSR3 as well as 3 other mutants (NDSR6, 11 and 15) were found in a distant cluster, although the latter three mutants had not previously been shown to lack markers on 7DL (see chapter two). Compared to the main group of NDSR, these four mutants were different for more than half of the SNPs assayed. It was very unlikely that these two groups only share half of the genetic background. However, results usually followed the same trend, hence the good correlations. In addition, the signal differences (i.e. differences between θ or R values) between these two groups were small, suggesting that most of these SNPs were false positives. Because the detection of polymorphisms between the NDSR mutants relied only on one replicate, the high number of SNPs detected between the group NDSR3/6/11/15 and the main one was most probably due to a technical issue rather than a biological one. Under those circumstances, this cluster was removed from the analysis. The detection of loci

on 7DL only relied on the difference between NDSR12 and the reference group (NDSR01, NDSR02, NDSR04, NDSR05, NDSR07, NDSR10, NDSR13).

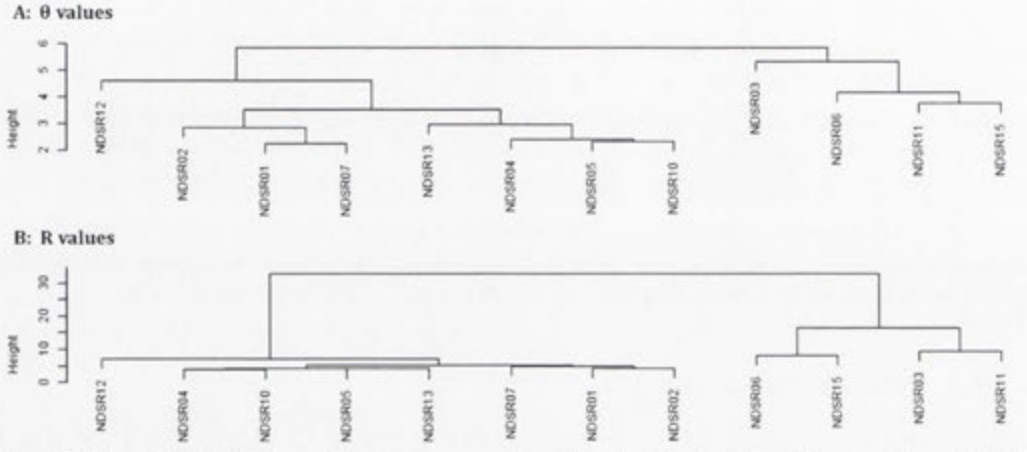


Figure 9: Hierarchical clustering analysis. Trees were constructed using the Euclidian distance matrices of NDSR mutants for A) the θ values and B) the R values. For each tree, two main NDSR clusters are visible: the reference cluster (01, 02, 04, 05, 07, 10, 13) and a distant cluster (03, 06, 11, 15). NDSR12, which lacks 7DL, is part of the main cluster but is the most dissimilar, as expected.

Any difference between the signals for NDSR12 and for the reference group could be interpreted as a missing locus on 7DL. Similarly than for the detection of polymorphism between Col and Col-NS766, loci on 7DL were detected through shifts in θ or R values. Out of the 8,632 SNPs assayed, 478 loci were considered to be putatively on 7DL. Most of them had different R values between NDSR12 and the reference group (83%), suggesting a missing locus. This was in agreement with a missing chromosome arm. However, not many loci were detected in single copy (present in the reference group and absent in NDSR12). Most of these loci might be present on several chromosomes, with at least one copy on 7DL (Figure 7D).

I.3. Selection for mapping

35 loci were candidates for mapping, as they were polymorphic, between Col and Col-NS766, and inferred to be on 7DL. Their chromosomal locations were checked by aligning (BLASTN) the sequences used for designing the chip probes, against the wheat GSS database, which contains sequences known to originate from specific wheat chromosome arms. Although always found on 7DL, the selected loci were also present on the homoeologous genomes, with 7AL being often the first hit. Among the 35 candidate loci, the most promising 24 were selected for mapping.

I.4. Genetic mapping

With only two not showing any difference between Col and Col-NS766, possibly due to primer issues, all markers could distinguish the two lines. This showed the 9k chip reliability for genotyping and finding new polymorphisms. In order to quickly test gene linkage, selected markers were first tested on a subset of 45 homogeneous F_2 plants (22 susceptible and 23 resistant). This permitted rapid determination of those that are the most closely linked to the locus, that is, those with the least recombinants between the marker and the stem rust resistance. Among the 22 markers tested, 17 showed association with the phenotype, of which four were completely associated (c31955, c6142, c916, c24796), based on the population subset. These four markers were tested on an additional set of 109 F_2 plants. Unfortunately, none showed better linkage than the previous SSRs (wmc0273 and cfa2040) but two, c31955 and c916, were co-segregating with these SSRs. These co-segregating markers were separated from the rust resistance by 10 recombination events in the $F_{2:3}$ population comprising 154 families.

Despite the relatively high number of loci found polymorphic between Col and Col-NS766, the use of the 9k chip was unsuccessful in obtaining markers that were closer to the gene of interest. More markers could have been tested but as their chromosomal location was unsure, they would have been selected randomly without the certainty of a better linkage. For this reason, the GBS approach was undertaken.

II. Genotyping by sequencing

DNA pools from stem rust-resistant and susceptible bulks of F_2 plants were prepared based on the evaluation of $F_{2:3}$ genotypes with race #313 and Ug99 in 2012 (see chapter two). The resistant bulk (BR) was composed of 36 homozygous resistant individuals and the susceptible bulk (BS) of 34 homozygous susceptible. For each bulk, as well as for the parental lines, Col and Col-NS766, two libraries were made using the methylation-sensitive restriction enzymes, *AatII* or *PstI*. All libraries were sequenced on Illumina® HiSeq 2000 for 100bp paired-end reads.

II.1. Detection of linked markers

After quality trimming and debarcoding, 37 million paired-end reads, in average, were generated per library. Reads were aligned against the whole GSS database, and 79% of them aligned to 1.4 million contigs, in average, per library. As expected, PE1 reads aligned to more contigs than PE2 did, due to the variability in the region sequenced. 865,000 SNPs, in average, were found between each library and the GSS sequences. After comparing each SNP between the corresponding libraries, 378 and 588 SNPs were classified as GSP 1 and 2 (Table 1) in the *AatII* and *PstI* libraries, respectively. These SNPs were contained in 242 and 378 GSS contigs, respectively. Surprisingly, most of these SNPs were located on contigs from chromosome 7AL, the others being evenly spread throughout the genome (Figure 10).

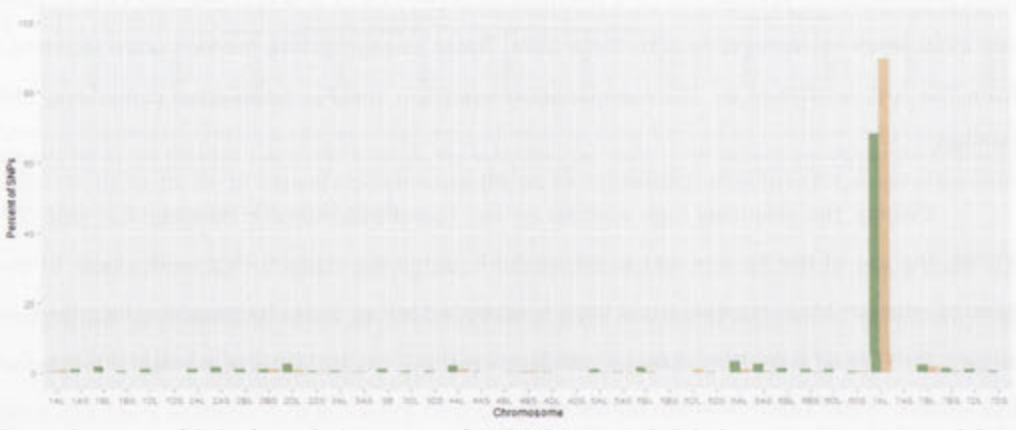


Figure 10: Percent of SNPs, for each chromosome, found to be putatively linked to stem rust resistance. Colours represent the enzyme used for the libraries construction: *AatII* (green) and *PstI* (orange).

II.2. Genetic mapping

Among the 966 SNPs putatively linked to stem rust resistance, 32 were selected for genetic mapping, based on various criteria such as only GSP 1, one SNP per contig, contigs longer than 5 kbp, no repetitive regions, synteny with Rice and *Brachypodium*... Similarly to the 9k chip experiment, markers were first tested on a population subset before the full population (max 196 F_2 lines). In addition, as so many SNPs were found on chromosome 7AL, when they were expected to be on 7DL, markers were also mapped to group 7 chromosomes using Chinese Spring (CS) and three CS nullisomic lines, each missing one of the group 7 chromosomes: 7A (CS-N7A), 7B (CS-N7B) and 7D (CS-N7D). As shown in Figure 11, markers being linked to the segregating locus usually displayed a different signal with CS-N7A than with CS and the other deletion lines. Most of the time, the signal with CS-N7A and the non-template

control (NTC) were similar, implying that no amplification occurred when the chromosome 7A was removed. This result confirmed that the segregating locus was located on 7A in the Col x Col-NS766 population, and not on 7D as it was thought. This unexpected finding dramatically changed the project focus away from the 7DL region, to further characterization of this locus on 7AL.

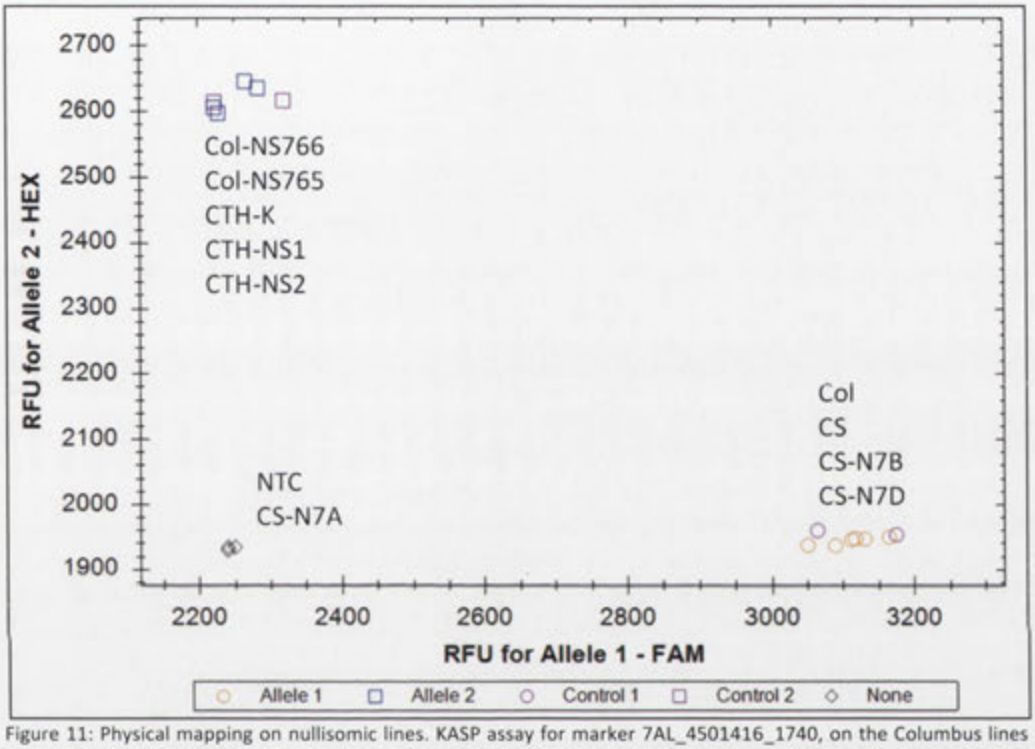


Figure 11: Physical mapping on nullisomic lines. KASP assay for marker 7AL_4501416_1740, on the Columbus lines (Col, Col-NS765, Col-NS766), the Canthatch lines (CTH-K, CTH-NS1, CTH-NS2), Chinese Spring (CS), Chinese Spring deletion lines (CS-N7A, CS-N7B, CS-N7D) and the non-template control (NTC).

Most markers could easily differentiate the alternative alleles, and a genetic map of 13.2 cM was drawn (Figure 12). A large number of markers co-segregated with each other, as well as with the previous best linked SSRs and SNP markers from the 9k chip (wmc0273, cfa2040, c31955 and c916). Although the use of a population subset may have led to a bias in the selection of markers for full population mapping, the high density of co-segregating markers indicated that this region was highly likely to be closely linked to the resistance locus, as per the principle of BSA. Nevertheless, just like with the previous markers, there were still several recombinants between these markers and the phenotype (Table 2).

However, these recombinants were not conformed to a mapping population segregating for one gene, and the gene of interest could not be located. Indeed, for ten of these

families (A03, B16, B42, A51, B01, B02, B03, B07, B15 and B64), all markers were either coming from the resistant parent (homozygote R) or from both parents (heterozygote H), but the corresponding phenotypes were segregating (H) and susceptible (S), respectively (Table 2 left). Some markers were also tested on additional families descended form the F_1 plant C, and showed similar results than with plants descended from F_1 plants A and B. To make sure these families were properly assessed, 12 F_3 plants from each family were individually phenotyped with race #313 and genotyped with 7AL_4501416_1740. Results confirmed that the genotypes and phenotypes shown for these families were correct.

The results for the non-conforming $F_{2:3}$ families indicated that the gene of interest was located away from the cluster of co-segregating markers. However, there were more recombination events between the phenotype and markers on both the distal and proximal sides of the cluster, indicating that the cluster represented the region on chromosome 7AL that was most closely linked with the phenotype. These data indicate that stem rust resistance in Col-NS766 may not be attributable to a single genetic locus.

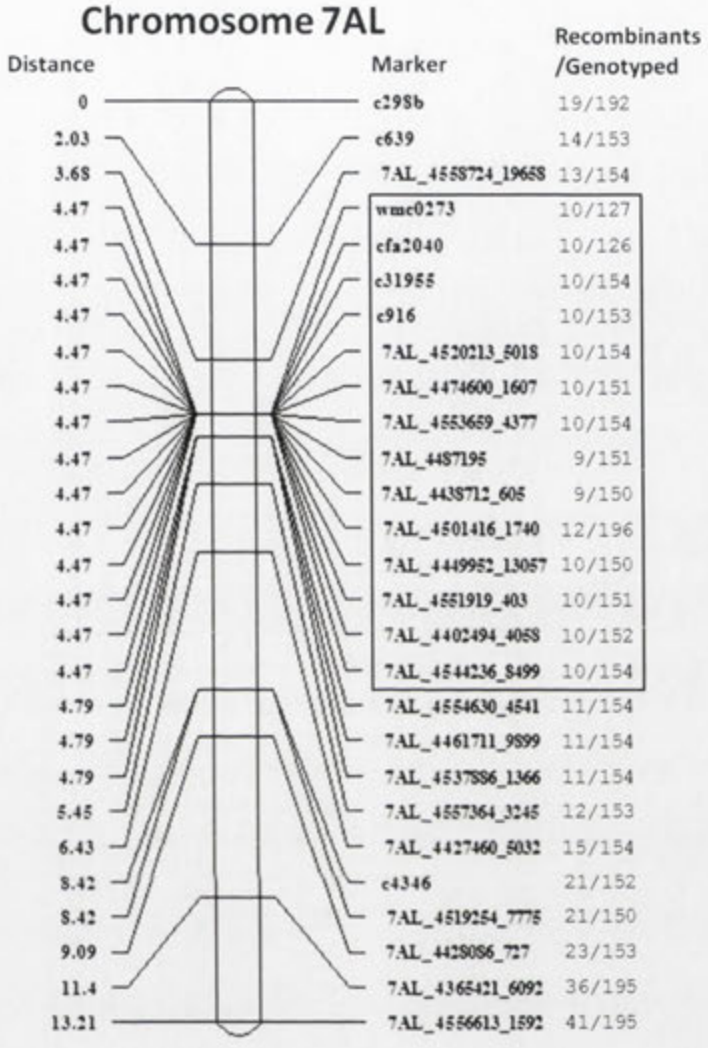


Figure 12: Genetic map of the 7AL resistance locus. Distances are in Kosambi cM and relative to the first marker at the top, c298b. The number of recombinants and F_2 plants genotyped are indicated on the right. The cluster of co-segregating markers is enclosed in the rectangle.

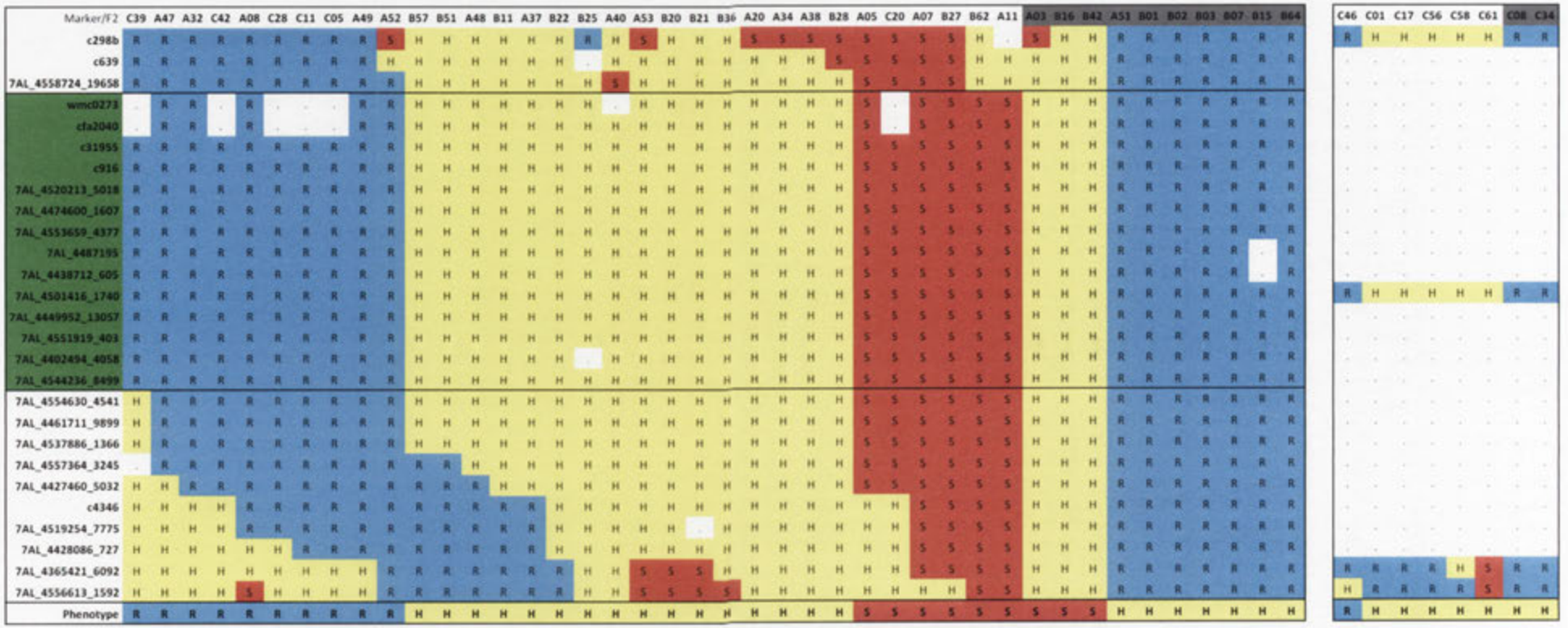


Table 2: Genotypes and phenotypes of recombinants in the Col X Col-NS766 $F_{2:3}$ mapping population. Genotypes of all F_2 plants for which at least one marker showed recombination with the phenotype (determined in F_3 progeny). The left part of the table represents families that were genotyped for all markers, and the right part represents extra families that were genotyped for few markers. The cluster of co-segregating markers is indicated in green, and the non-conforming families in grey. Letters represent the F_2 allele or the $F_{2:3}$ phenotype: homozygous for the Col-NS766 allele or resistant (R); heterozygote or segregating (H); homozygous for the Col allele or susceptible (S); missing (.).

DISCUSSION

Comparison of genotyping techniques

Both the 9k chip and the GBS are powerful techniques that enable rapid detection of polymorphism. With only one run, both methods provided many markers, which identified a locus on 7AL that was linked to the rust resistance phenotype. The proportion of mapped markers that were linked was higher with GBS, but the selection of markers with the 9k chip would probably have been better if coupled to BSA. The main advantages of the 9k chip lie in the short time necessary (4 days), the relative low cost and particularly the ease to analyse it (standard desktop computer). It is ideal for genome wide genotyping. Moreover, most SNPs present on the chip have now been mapped (Cavanagh et al., 2013) and the high density genetic map available makes this technology even more valuable. However SNPs are fixed and not necessarily relevant to the variety of interest, whereas it is not the case with GBS. The latter technique also gives a greater output, but is much harder to interpret and requires good computing resources. It is also more time demanding and costs more (although it depends a lot on the level of multiplexing). Finally, markers found with the 9k chip often detected homoeologous sequences, whereas those found through GBS were often genome specific. That said, it is not a real issue for genotyping when using KASP assays, as genome specific markers are not necessary for differentiating each zygosity state, contrary to older methods (e.g. CAPS).

Resistance locus on chromosome 7AL

At first, efforts were put into finding markers that were located on 7DL in order to map the inactivated stem rust 7DL-Sup generated in the Canthatch mutants. However, through GBS, it became obvious that the locus being mapped and associated with stem rust resistance, in the Col x Col-NS766 population, was on 7AL. This emphasized the power of this method for finding linked markers, even without preliminary mapping knowledge. In retrospect, all linked markers found with the 9k chip were also on 7AL, according to the chip genetic map (Cavanagh et al., 2013) and the GSS database. Although the experiment was designed and analysed for finding polymorphism on 7DL, linked markers on 7AL were still found due to the presence of homoeologous alleles.

Concerning the previously mapped SSR markers (wmc0273 and cfa2040), because they were found to co-segregate with the markers located on 7AL, they should also be on this chromosome. This was suggested with tests using NDSR12 (lacking 7DL), which showed that the polymorphism between Col and Col-NS766 was not on 7DL (Figure Sup. 2). This also explained the low number of SSRs found to be polymorphic, as they were selected for being on 7DL and not 7AL. Just like the markers from the 9k chip, wmc0273 and cfa2040 were found linked because homoeologues were also present on 7AL and 7BL (Somers et al., 2004). According to the SSR and 9k chip maps (Cavanagh et al., 2013; Somers et al., 2004), the 7AL locus was located near the telomere.

Col-NS766 was developed through the backcrossing of one of the Canthatch (CTH)-NS mutants with Col (Kerber, 1991). As expected, the 7AL locus in Col-NS766 was inherited from the CTH background, as all CTH lines showed the same allele for all linked markers (Figure 11). Furthermore, tests of 8 markers, including some from the cluster of co-segregating markers, suggested that the 7AL locus in Col-NS766 and CTH, originated in Thatcher, as all loci showed the same allele (data not shown). Moreover, the independent backcrossed line Col x CTH-NS (independent CTH-NS than for Col-NS766), Col-NS765, also had this locus introgressed, supporting that it is necessary for the resistance against stem rust.

Col may lack the 7DL-Sup

Like CTH, Col is related to Thatcher, which is thought to be the source of the 7DL-Sup. Thus, Col was thought to possess the 7DL-Sup, as well as the masked resistance genes. It was believed that the resistance in Col-NS was due to the introgression of the NS allele from CTH-NS into Col, allowing the expression of suppressed resistance gene(s) in the Col genetic background. Although it was reported that the 7DL-Sup is not uncommon in wheat, as well as in *Ae. tauschii* (Kerber, 1983), it is worth considering the hypothesis that Col does not possess the 7DL-Sup, nor the resistance gene(s) suppressed by it. In this case, introgressing the NS allele in Col would not confer resistance, but the introgression of the suppressed resistance gene(s) would. This hypothesis is consistent with the mapping data, which strongly suggested that the resistance in Col-NS was not due to the introgression of the NS allele on 7DL, but of a resistance locus on 7AL.

Presence of the 7DL-Sup in Col (susceptible) was tested by crossing Col with CTH nullisomic 7D (CTH-N7D, resistant) or ditelosomic 7DS (CTH-Dt7DS, resistant). F_2 progeny were

phenotyped with stem rust (race #313). If Col had the 7DL-Sup, all F_2 plants with at least one copy of the 7DL chromosome would be susceptible. On the other hand, if the 7DL-Sup was absent in Col, the 7DL chromosome would not affect the resistance. Unfortunately, resistant and susceptible phenotypes were too close to each other and it was not possible to tell them apart. Even the resistant positive controls CTH-N7D and CTH-Dt7DS looked very much like the susceptible CTH-K. This experiment was not repeated as race #313 was not appropriate for accurately differentiating plants with CTH background. Tests will need to be undertaken with another stem rust race, such as Ug99, in order to test for the presence of the 7DL-Sup in Col.

The mapping of a locus linked to stem rust resistance, on chromosome 7AL, significantly changed the focus of the project. However, the discovery of a suppressed resistance gene would be an important contribution to understanding the mode of action of the 7DL-Sup. Moreover, removing the 7DL-Sup would only be useful if the resistance genes are present. Developing robust markers for the 7AL resistance locus would greatly help for its introgression in marker-assisted breeding programs. Furthermore, cloning the suppressed resistance gene might help in finding the 7DL-Sup. For instance, both genes could be homoeologous genes, as in the case of *Pm8* and *Pm3*. *Pm8*, which confers resistance to powdery mildew and was introgressed in wheat from rye, was found suppressed by the orthologous wheat gene *Pm3* (Hurni et al., 2013; McIntosh et al., 2011).

Resistance locus on chromosome 6A

Although not available at the time of this experiment, most SNPs present on the 9k chip were subsequently mapped, and chromosome locations for most loci became available (Cavanagh et al., 2013). Concerning the loci found polymorphic between Col and Col-NS766, very few were located on 7DL (4%) and more were on 7AL (15%). However, more than half of the polymorphic loci (55%) were located on chromosome 6A. This was highly surprising because after five backcrosses with Col, Col-NS766 was not expected to possess many regions introgressed from CTH other than the one involved in the resistance; particularly such a large region, which contained most of the polymorphic loci. Most polymorphic loci on 6A were found to have different θ values in Col and Col-NS766, thus implying that this region was present in both lines and that polymorphism was not due to a large deletion. Two markers of 6A were developed and mapped in the mapping population. They were found linked between each other and at a genetic distance of 5.1 cM. No linkage was found with the phenotype.

Genome	A		B		D	
Chr.	#	%	#	%	#	%
1	2	1%	8	3%	2	1%
2	6	2%	10	4%	2	1%
3	0	0%	0	0%	0	0%
4	12	5%	0	0%	0	0%
5	2	1%	10	4%	0	0%
6	137	55%	8	3%	1	0%
7	37	15%	17	7%	9	4%

Table 3: Chromosomal distribution of the 9k chip loci that were polymorphic between Col and Col-NS766. The majority of polymorphic loci were found on chromosome 6A, and only 15% were found on 7A.

For some reason, perhaps only by chance, this large fragment was selected in Col-NS766. Nevertheless, some components present on 6A might affect the resistance and, although not essential, were selected in Col-NS766 for full resistance. Interestingly, the two markers of 6A showed that Col-NS765 carried the same alleles than Col-NS766. Moreover, transcriptome analysis showed that many genes on 6A were differentially expressed between Col-NS765/Col-NS766 and Col, most likely due to a large polymorphic region (see chapter four). Because Col-NS765 and Col-NS766 are two independent backcrossed lines in Col, this suggests that the 6A locus could have been actively selected, and thus could affect the phenotype.

However, this does not explain the large size of this introgressed fragment. One possibility is that this selected region does not recombine much with Col, such as the centromere. This was suggested by the 9k chip map, on which most polymorphic loci on 6A were located around the centromere and delimited to a 22 cM wide region. Another possibility is that two or more genes influencing the phenotype are linked to each other on chromosome 6A, but are still apart from each other. Because the chance of double recombination between them remains minimal, the selection of these genes during backcrosses would also have lead to the selection of the region between them.

Multigenic resistance

The phenotypic data were consistent with segregation of a single gene in the mapping population and mapping showed that it was located on chromsome 7AL. However, the fine mapping failed to place the gene on a genetic map and it is unlikely that it is located outside the map. Indeed, the centre of the map, especially at the cluster of co-segregating markers

level, showed the least recombinants. Therefore, these markers should be the closest to the gene of interest, since the further from the gene, the more the recombinants. In addition, it is unlikely that all markers, notably those found through BSA, mapped a region far from the gene. Under those circumstances, it is possible that the resistance involves multiple genes. This would not be surprising as Williams et al. (1992) suggested that the 7DL-Sup may inhibit the expression of as many as three or more genes for resistance in CTH. In addition, this would be in agreement with phenotypic observations of individual F_3 plants, suggesting one essential gene and other optional genes (see chapter two).

One gene *A* would co-segregate with the cluster of markers on 7AL, as the genetic map indicates that this region is the most suited for containing a resistance gene. This gene explains most phenotypes in the $F_{2:3}$ families, and thus would be essential for the resistance. Moreover, only two types of non-conforming families were found: i) segregating families that carried only the resistant allele and ii) susceptible families that were heterozygous.

Resistant F_3 families were always derived from F_2 individuals homozygous for the 7AL locus and F_3 families that gave segregating phenotypes were always derived from F_2 individuals that contained at least one copy of the 7AL locus. This is consistent with the locus on 7AL being necessary but not sufficient for full resistance. The other gene(s) could be anywhere else on the genome, but some conditions would be required in order to still retain a segregation ratio that resembles that of one-gene segregation (i.e. 1:2:1). Several models are possible.

Speculative models

The first one, m1, involves a gene *B* that is on 7AL and linked to the main gene *A*. Both genes would most of the time co-segregate, and thus would be considered as one locus. In some cases, a recombination event between the two genes would occur, explaining the unexpected phenotypes in the non-conforming families. With these genes being 8 cM apart (12 recombinants out of 158 $F_{2:3}$ families), the observed results would fit the expected ones (Table 4). However, it is most likely that the second gene is located somewhere else on the genome, notably on chromosome 6A. Indeed, results from the 9k chip indicated a large introgression on 6A in Col-NS766. This was not found with RAD-Seq due to the BSA, which focused on the main gene on 7AL. The 6A fragment seems to be also present in Col-NS765 and may have been selected because it contains genes that affect the resistance. While models

with two unlinked genes (on 7AL and on 6A) do not fit the observations, models with 3 genes do.

One model, m2, involves the main gene *A* and two unlinked genes *B* and *C*. Of these last two, only one is necessary for resistance, that is, they both provide similar resistance when coupled to *A*, but they are not both required to be present for good resistance. *B* and *C* would have been selected because they may have additive effects and provide a higher resistance when present together. However in this model, too many susceptible families that only possess the resistant allele of *A* are expected, when none was observed.

Another model, m3, involves the main gene *A* and two linked genes *B* and *C* that are located at 22 cM from each other. This distance was based on the possible size of the 6A fragment, according to the 9k chip map. *B* and *C* would be essential for resistance, but would in some manner interact with each other. However, they would only be effective if they both come from the same parent (i.e. B and C from one parent, or b and c from the other). A recombination event between these two genes would result in susceptibility, as *B* would come from one of the parent and *C* from the other. Alleles from CTH-NS would have been selected in Col-NS766 because they provide better resistance than that of Col. The expected frequency of resistant families is lower than the observed one, but the latter may have been biased due to the rejection of families that did not give consistent phenotypes across rust races or years. Indeed, resistant families were easier to score, as resistant F_3 plants were less variable. Therefore, more resistant families may have been scored compared to segregating and susceptible families, for which the phenotypes in F_3 were too variable to accurately score them.

Finally, a similar model, m4, also involves the main gene *A* and two linked genes *B* and *C* that are located at 22 cM from each other. However, *B* provides some degree of resistance and is epistatic to *C*, which completes the resistance. From all models, this one gives expected phenotype frequencies that fit the best the observed ones, but the frequency of susceptible non-conforming families is low.

Although these four models are possible, they all have weaknesses that cannot be resolved with the present data, as too few $F_{2:3}$ families were tested to confidently determine if they are correct. Testing more families will be needed, especially if susceptible families carrying only the resistant allele for the main gene on 7AL were observed, as these families are

expected in any model. However, it is important first, to confirm that the gene on 7AL truly co-segregates with the cluster of markers. To do this, populations that only segregate for this gene must be developed and mapped. This could be done using some of the $F_{2:3}$ families that segregate for the resistance and were found to be heterozygous for the cluster of co-segregating markers. Assuming that there is no susceptible allele of the 7AL resistance gene in Col, families that presented three times more resistant F_3 plants than susceptible (i.e. ratio 3:1) may be suitable for quickly developing such a population, as the $F_{3:4}$ families should only segregate for the 7AL gene. The A44 family fits this ratio.

Model			m1	m2	m3	m4	Observations
Number of genes		1 gene	2 genes	3 genes	3 genes	3 genes	
Gene linkage		-	A and B are linked (8 cM)	All genes are unlinked	B and C are linked (22 cM)	B and C are linked (22 cM)	
Phenotype conditions	R	AA	AA + B	AA + B or C	AA + BC or bc	AA + BC or bc (AAB: partial R)	
	S	aa	aa ot bb	aa or bbcc	aa or BBcc/bbCC	aa or bbCC	
Phenotype frequency	R	0.25	0.21	0.17	0.16	0.19	0.21
	I	0.50	0.50	0.53	0.57	0.56	0.54
	S	0.25	0.29	0.30	0.27	0.26	0.25
Phenotype number	R	39.5	33.4	27.2	25.5	29.2	33
	I	79.0	79.0	83.9	90.4	88.0	86
	S	39.5	45.6	46.9	42.1	40.8	39
Chi-square test	χ2	1.696	1.572	2.641	2.631	0.608	
	p-value	0.4282	0.4556	0.2609	0.3389	0.7566	
Frequency of non-conforming families	AA: I	0.000	0.037	0.063	0.083	0.062	0.057
	Aa: S	0.000	0.037	0.031	0.011	0.006	0.019
	AA: S	0.000	0.002	0.016	0.006	0.003	0.000
Weakness		No non-conforming families	Same frequency of non-conforming families	1.6% of susceptible families with R allele only	Low frequency of resistant families	Low frequency of susceptible non-conforming families	

Table 4: Multigenic resistance models. Depending on the number of genes, their linkage and their effect on stem rust resistance, the expected phenotypes for $F_{2:3}$ families were calculated. The phenotype conditions indicate which allele is necessary for the families to be scored as resistant (R) or susceptible (S). Families are always considered segregating (I) if resistant and susceptible F_3 plants are expected. If none of these phenotypes are expected in F_3, families are considered resistant. The frequency of non-conforming families represent the families with phenotypes that do not correspond to the expected one, assuming the allele for the essential gene on 7AL. Models were approved using chi-square goodness-of-fit tests between expected and observed phenotypes, as well as on the frequency of expected non-conforming families.

MATERIALS AND METHODS

Identification of SNPs using wheat 9k SNP chip

DNA samples of Col, Col-NS766 and NDSR mutants were genotyped for the 8,632 gene-based SNPs using the wheat 9k Infinium™ iSelect BeadChip assay (Cavanagh et al., 2013) on the Illumina® iScan instrument following the manufacturer's instructions. Genotypic analysis was performed using R (http://www.R-project.org) and loci were considered present

for signal intensity R ≥ 0.14. Polymorphism was found by comparing the normalised θ and R values. Values were considered significantly different if the confidence intervals (95%) did not overlap and if there was a difference between the means of at least 0.05 (θ only) or 10% of the max R mean (R only). SNPs were selected for genetic mapping if polymorphic between Col and Col-NS766 (based on θ values) and between NDSR12 and the reference NDSR group.

Identification of linked SNPs using GBS

70 F_2 comprising 36 homozygous resistant and 34 homozygous lines (based on F_3 families scoring with race #313 and Ug99 in 2012) were selected to compose the two bulks. The genome complexity of Col, Col-NS766 and the bulks was reduced through the construction of RAD tag libraries (Baird et al., 2008) using methylation-sensitive restriction enzymes *AatII* and *PstI* and random fragmentation with Covaris® S220 Focused-ultrasonicators. Fragments of 450 bp approximately (excluding adapters) were selected on 1.5% agarose gel and purified using QIAquick Gel Extraction Kit (QIAGEN®). Libraries quality was checked on Agilent® 2100 Bioanalyzer and quantified using KAPA Library Quantification Kit (KAPABIOSYSTEM®). Libraries were sequenced on Illumina® HiSeq 2000 instrument for 100 bp paired-end, using 4 libraries per sequencing lane. Sequences were quality filtered using custom perl scripts and PE2 reads (corresponding to the restriction site end) were adapter trimmed to remove an internal barcode introduced during library preparation, which reduced their length to 95 (*AatII*) and 97 (*PstI*) bp. Reads over 50bp were individually aligned against all chromosome contigs from the wheat GSS (http://www.wheatgenome.org/) using BWA v0.5.9-r16 (Li and Durbin, 2009) with 2 mismatches allowed. Pileup files were then generated using SAMtools v0.1.14 (Li et al., 2009) and SNPs between each library and the reference were discovered using custom perl scripts. Alignment of at least 7 reads at a sequence variant position was required to call a genotype, where homozygous SNP genotypes had an alternate allele frequency < 5%. SNPs were considered putatively linked to the gene of interest if polymorphic between Col and Col-NS766 and between the bulks, and in phase between the bulks and the appropriate parent. Only SNPs on homozygous genotypes (GSP) were selected for genetic mapping.

Genetic mapping

The 9k chip reference sequences and the GSS contigs were directly used for marker development. KASP markers were designed and assayed following manufacturer's protocol (KBioScience, UK). Following PCR, end-point fluorescence was measured using BioRad® CFX96

real-time instrument and genotypes were determined using CFX Manager v3.2. The genetic map was built using QTL IciMapping v3.2 (Wang et al., 2012) and distances were calculated using the Kosambi mapping function.

REFERENCES

Allen, A.M., Barker, G.L.A., Berry, S.T., Coghill, J.A., Gwilliam, R., Kirby, S., Robinson, P., Brenchley, R.C., D'Amore, R., McKenzie, N., et al. (2011). Transcript-specific, single-nucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (*Triticum aestivum* L.): Wheat SNP markers. Plant Biotechnol. J. *9*, 1086–1099.

Asad, M.A., Xia, X., Wang, C., and He, Z. (2012). Molecular mapping of stripe rust resistance gene *YrSN104* in Chinese wheat line Shaannong 104. Hereditas *149*, 146–152.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., and Johnson, E.A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE *3*, e3376.

Bennetzen, J.L., Schrick, K., Springer, P.S., Brown, W.E., and SanMiguel, P. (1994). Active maize genes are unmodified and flanked by diverse classes of modified, highly repetitive DNA. Genome *37*, 565–576.

Catchen, J.M., Amores, A., Hohenlohe, P., Cresko, W., Postlethwait, J.H., and De Koning, D.-J. (2011). Stacks: building and genotyping loci de novo from short-read sequences. Genes Genomes Genetics *1*, 171–182.

Cavanagh, C.R., Chao, S., Wang, S., Huang, B.E., Stephen, S., Kiani, S., Forrest, K., Saintenac, C., Brown-Guedira, G.L., and Akhunova, A. (2013). Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. Proc. Natl. Acad. Sci. *110*, 8057–8062.

Davey, J.W., and Blaxter, M.L. (2011). RADSeq: next-generation population genetics. Brief. Funct. Genomics *9*, 416–423.

Fellers, J.P. (2008). Genome Filtering Using Methylation- Sensitive Restriction Enzymes with Six Base Pair Recognition Sites. Plant Genome J. *1*, 146.

Henry, R.J., Edwards, M., Waters, D.L.E., S, G.K., Bundock, P., Sexton, T.R., Masouleh, A.K., Nock, C.J., and Pattemore, J. (2012). Application of large-scale sequencing to marker discovery in plants. J. Biosci. *37*, 829–841.

Hurni, S., Brunner, S., Buchmann, G., Herren, G., Jordan, T., Krukowski, P., Wicker, T., Yahiaoui, N., Mago, R., and Keller, B. (2013). Rye *Pm8* and wheat *Pm3* are orthologous genes and show evolutionary conservation of resistance function against powdery mildew. Plant J. *76*, 957–969.

Kerber, E.R. (1983). Suppression of Rust Resistance in Amphiploids of *Triticum*. In Proc. 6th International Wheat Genetics Symposium, (Kyoto, Japan), pp. 813–817.

Kerber, E.R. (1991). Stem-rust resistance in "Canthatch" hexaploid wheat induced by a nonsuppressor mutation on chromosome 7DL. Genome *34*, 935–939.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Liu, S., Yeh, C.-T., Tang, H.M., Nettleton, D., and Schnable, P.S. (2012). Gene mapping via bulked segregant RNA-Seq (BSR-Seq). PloS One *7*, e36406.

Ma, H., Kong, Z., Fu, B., Li, N., Zhang, L., Jia, H., and Ma, Z. (2011). Identification and mapping of a new powdery mildew resistance gene on chromosome 6D of common wheat. Theor. Appl. Genet. *123*, 1099–1106.

McIntosh, R.A., Zhang, P., Cowger, C., Parks, R., Lagudah, E.S., and Hoxha, S. (2011). Rye-derived powdery mildew resistance gene *Pm8* in wheat is suppressed by the *Pm3* locus. Theor. Appl. Genet. *123*, 359–367.

Michelmore, R., Paran, I., and Kesseli, R.V. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. *88*, 9828–9832.

Poland, J.A., and Rife, T.W. (2012). Genotyping-by-sequencing for plant breeding and genetics. Plant Genome J. *5*, 92.

Poland, J.A., Brown, P.J., Sorrells, M.E., and Jannink, J.-L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS ONE *7*, e32253.

Somers, D.J., Isaac, P., and Edwards, K. (2004). A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). Theor. Appl. Genet. *109*, 1105–1114.

Trick, M., Adamski, N.M., Mugford, S.G., Jiang, C.C., Febrer, M., and Uauy, C. (2012). Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploid wheat. BMC Plant Biol. *12*, 14.

Wang, J., Li, H., Zhang, L., and Meng, L. (2012). Users' Manual of QTL IciMapping Version 3.2.

Williams, N.D., Miller, J.D., and Klindworth, D.L. (1992). Induced mutations of a genetic suppressor of resistance to wheat stem rust. Crop Sci. *32*, 612–616.

SUPPLEMENTAL

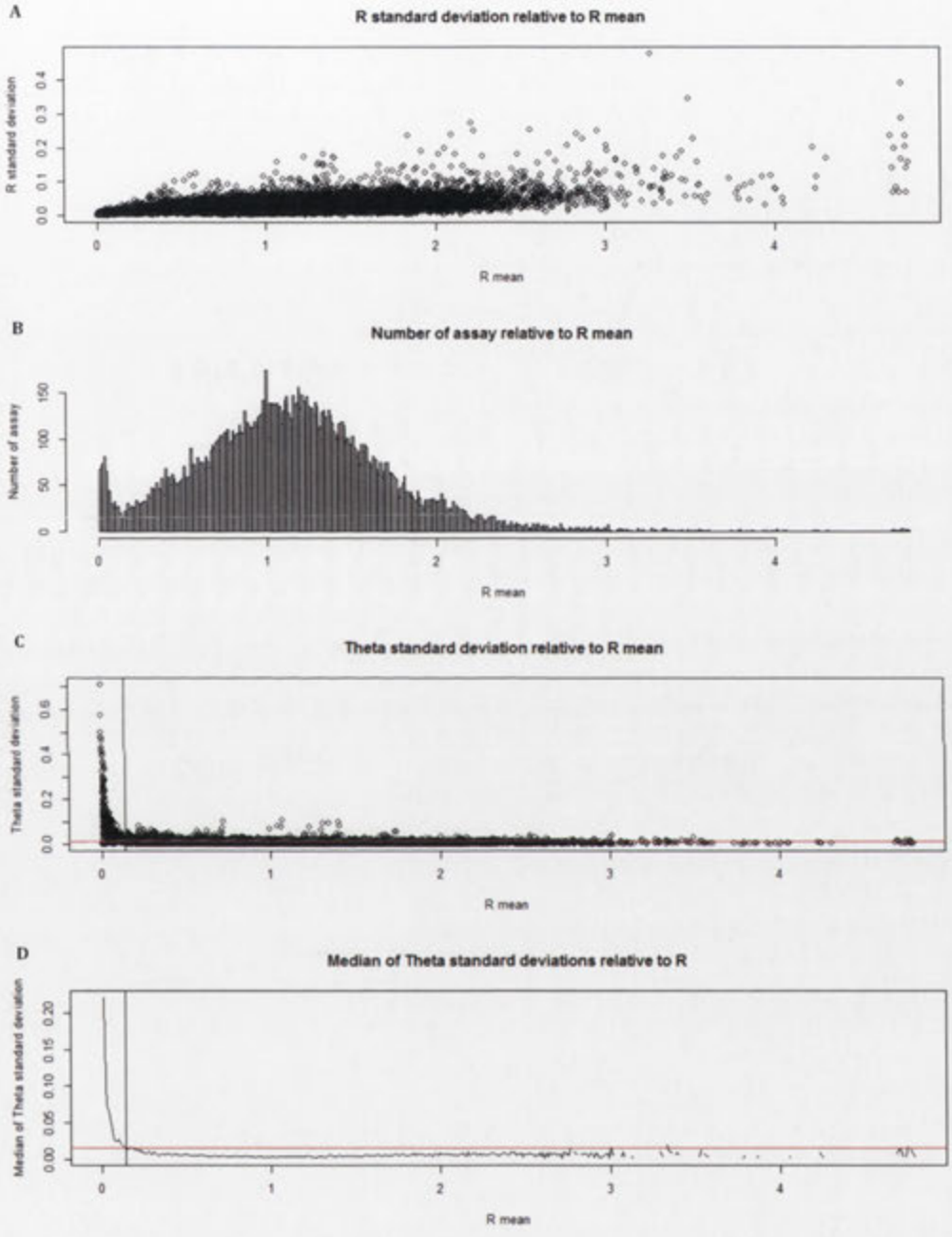


Figure Sup. 1: 9k chip statistics. A) Standard deviation of the signal intensity R in relation to the mean. The SD remains low at any intensity. B) Number of assay that gave a particular signal intensity R. Most assays had signal intensity between 0.5 and 1.5. C) Standard deviation of the normalised θ value in relation to the mean of the signal intensity R. D) Median of the standard deviation of the normalised θ value in relation to the mean of the signal intensity R. For both C and D, the SD is high for low signal intensities, whereas it is low for higher intensities. The horizontal red line represents the 9th percentile of the Theta standard deviations and the vertical black line represents the R threshold used for considering a locus present.

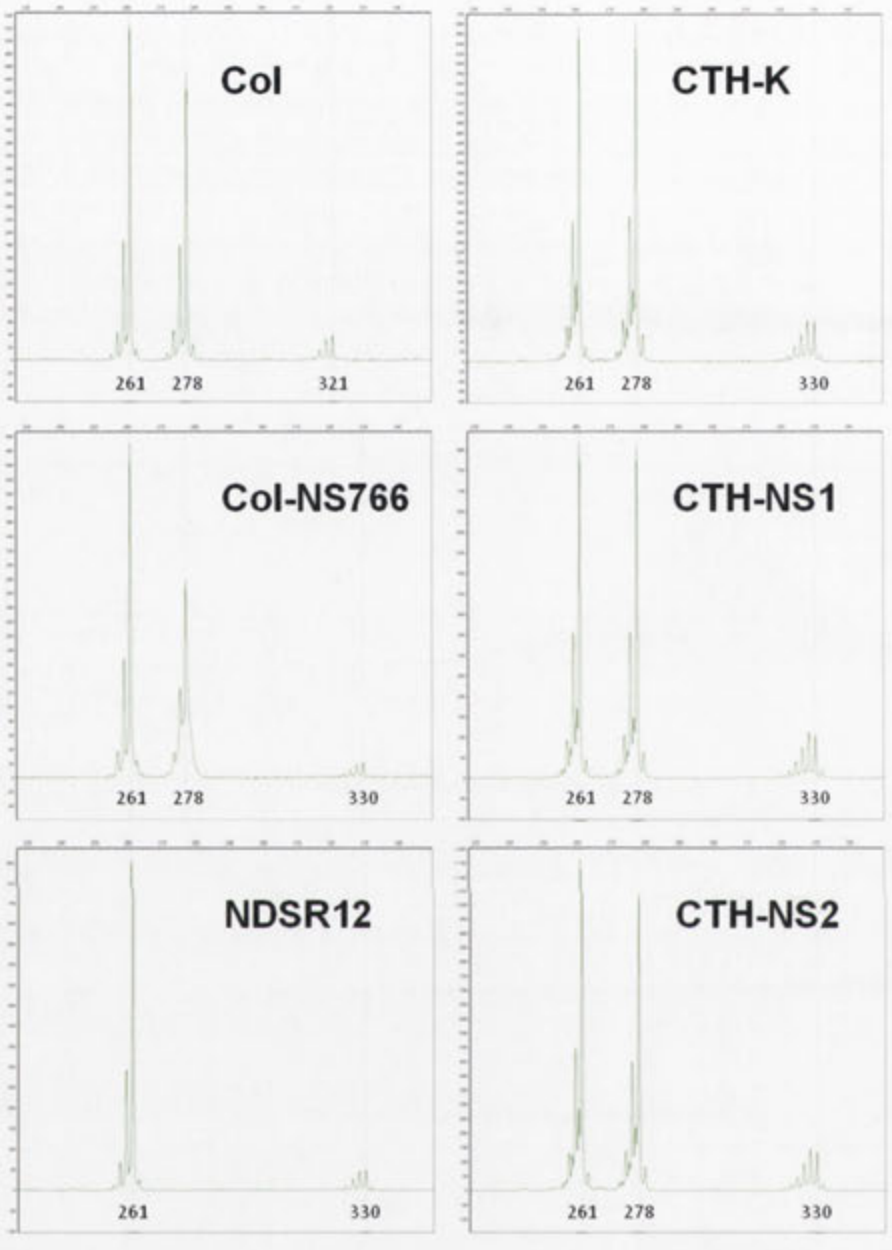


Figure Sup. 2: DNA fragment analyzer (DFA) output with SSR marker cfa2040. Each series of peaks represents an amplified allele and numbers represents their relative size. The polymorphic allele (321/330) is not on 7DL as it is present on NDSR12. The 7DL allele corresponds to the series of peak that is absent in NDSR12 (278).

GBS	Allele 1	Allele 2	Common
7AL_4365421_6092	AACTACCAGGAATCAGTTTCTAGTCACTTC	GAACTACCAGGAATCAGTTTCTAGTCACTTT	TGGTTCTTGGTGCACCCTAGTAGT
7AL_4402494_4058	TGAACTCACTTGAACAGCGCAA	TGAACTCACTTGAACAGCGCAT	GACGTGAAACGAATGGATCATGT
7AL_4427460_5032	CAGTTGTTACAGTATCAACACCTGGA	CAGTTGTTACAGTATCAACACCTGGG	TGTGGCGAATCAATCTAAGACGA
7AL_4428086_727	CAGTGAGGGAAACCCATATATTGAGT	AGTGAGGGAAACCCATATATTGAGC	GCAGCAATTTCGGATTGATGAGT
7AL_4438712_605	GCCTCGCCGACCCGTAAG	CCTCGCCGACCCGTCAA	CCATTGAGGTAATGGTGGAGGAG
7AL_4449952_13057	CACATGGATTACGAGGATCTTCAAC	ACACATGGATTACGAGGATCTTCAAT	CTTTGTGGTTGGATCTGATGACG
7AL_4461711_9899	CCGGCAAGGTCCGCTAGACT	CGGCAAGGTCCGCTAGACC	ACGAGATCCACATGGCTGCTT
7AL_4474600_1607	GCCAGATCATTTTCTGCTCTCG	GCCAGATCATTTTCTGCTCTCA	TAGGCAACTCTCTCGCCAAG
7AL_4487195	AGGGATCCTTGCAGAGAAGAGGTA	GATCCTTGCAGAGAAGAGGGG	CCGGCGGCTATCTAGAAATTCAT
7AL_4501416_1740	GTGAAGTGGGCTGCTAGCATTAAT	GTGAAGTGGGCTGCTAGCATTAAG	TCCGAATACGTCGCTCTATACGC
7AL_4519254_7775	CGCAACCTCAACACATGTACTCAG	CGCAACCTCAACACATGTACTCAC	AAGCCAGGGTATTCACCATACGA
7AL_4520213_5018	GCAAGCTTGAGTCCACATCCC	GCAAGCTTGAGTCCACATCCG	ACTGGATGACCCTTGATGAATCG
7AL_4537886_1366	GCGATGTTGAGGAACTACTCTGCT	CGATGTTGAGGAACTACTCTGCG	CAGGCATATCACACACACACAGC
7AL_4544236_8499	AAGGACAAGCGACCGCGC	AAAGGACAAGCGACCGCGT	TTCTTGACTACATGGCTGGTGGA
7AL_4551919_403	GATTATTCAAGTGCGACATCACCT	TTATTCAAGTGCGACATCACCA	TTGCCTAAGGCGGTATTTTG
7AL_4553659_4377	ATCATGAGGTTTGAGTTTGGCC	TCATGAGGTTTGAGTTTGGCG	CTAGGATGGACCCGTGGACAATA
7AL_4554630_4541	GCGCCACAGACCCTGTCG	GCGCCACAGACCCTGTCA	GCAAGAAACATCAGAAACTGCAGAA
7AL_4556613_1592	CATTTACAGGCGACCAGGCTAT	CATTTACAGGCGACCAGGCTAC	TGCACCTGCAGTATGAAAAACA
7AL_4557364_3245	CGCCCTTTCGAGACGCC	CCGCCCTTTCGAGACGCT	TGTTGCTACTGAGGCGGATT
7AL_4558724_19658	AACATGGAACCTGTCCCGG	GAACATGGAACCTGTCCCGA	AGTGGCATTTTCGTGTTGTCCTT

9k chip	Locus sequence
c298b	gttgatgcgctcactgatgctcgtgttacccatgaaatccagggcatctgggtctcatactggtcacagtgtgacgcgtccctcgggcagaagctggcat[T/C]GCGGCTCAAGATTAAGCCGAGCATGTAGATGGAAAAGGAGGA GGGAACCAGTCGCTCACCCGCCTCGGTTGGCGTCGATCGAGCCAAGGAGGGGTGCAGCA
c31955	ACACATACATGCATACATACATATGTATGGAGTCCACGACATCAGTCAATAAGTTACTCCTCCTGGGCCCATTATTATACCATATATCATATATGCACTA[T/C]TGGTGTTAAAATTAAAAATTCC TGCAGGGGASGAACAACAAGAGACATCTGCCTTTTCAGCAGCAGCACAACAAACGACACCTACCTATTCTTCCTCGGT
c4346	TTTGACGGCTTAGATTGGGTGACGCGGACGTGGGAGCTCTGTACACGGCAAGAGGCGTTTTTAAAGATCTCTTGTGCGCCCGGCCAGCAAGATGCAAATA[T/C]GTAAATGAATTAGTTACCC ATATACTGTAAATTAACGGCCAAGGAATATACCGCAGTAAAWTATGCAAGTAATAGTACTAGCTCYTTTTTTCTAACCTAT
c639	TACGGCCGGGGGAACTCGCCGAGGAGCGACGGCGCTAGGACCAACTCTAGGGGCAGATGATCATGATCAGGAGTGTCTATAGGATATGTAACCATGTAAC[T/C]GATATCCATTTGGAGAAT GTAAAGGTAGACCACCACACCATGCCATTGTTTCCCGTTGTCATGACGAAGTTTGCRCCGATCATGATGATCAATGTAATAG
c916	GCCTCCAACAAGTTAAAACATGCATTGATATCGGGTCAAACTGTGTTGAGATTGATAGGACAAAAAGACCTCGCATAGAAGAAATTGTTGACTCGCTCAA[T/C]GGACTACGTTCAAGCCAAA AAGTTACCGAGGTCTCCCGTCCGCCTACAACTGGCAAGAAGCTAATGGGCTCCATTTTTGGTCTGAAGAAATGAAACATGC

Table Sup. 1: List of SNP markers found linked to the stem rust resistance locus on 7AL. All markers were tested as KASP assays. Top panel) Markers found using genotyping by sequencing (GBS). "Allele 1 and 2" represent the primer sequences that are specific to one or the other allele, and "common" represents the third primer sequence in common to both reaction. Bottom panel) Markers found using the 9k SNP chip. Primer sequences are not available because they were designed by KASP manufacturer (http://www.lgcgenomics.com). Locus sequences from the chip are represented. Nucleotides in brackets represent the SNP.

Chapter 4

Characterization of the resistance conferred by the 7AL locus, through histological and transcriptome analyses

ABSTRACT

The resistance mechanism against wheat stem rust, conferred by the 7AL resistance locus in Columbus-NS765 and Columbus-NS766, was investigated. Histological observations first showed differences between resistant and susceptible lines at two days post-inoculation, during the penetration stage into the leaf through the stomatal opening. The resistance consisted in cell death, usually localised at the guard and epidermal cells, adjacent to an infection site. RNA-Seq of Columbus and the two resistant lines resulted in the detection of 353 genes differentially expressed and potentially involved in the resistance response. Several genes coding for cysteine-rich receptor-like kinases were notably found as good resistance genes candidates. Finally, data suggested the possible involvement of genes on chromosome 6A in the defence mechanism.

List of major abbreviations

7DL-Sup Canthatch 7DL suppressor; **Col** Columbus; **Col-NS** Columbus-NonSuppressor; **CRKs** cysteine-rich receptor-like kinases; **CTH** Canthatch; **CTH-NS** Canthatch-NonSuppressor; **DE** differentially expressed; **DPI** days post-inoculation; **ETI** effector-triggered immunity; **FC** fold change; **FDR** false discovery rate; **GSS** genome survey sequencing; **JA** jasmonic acid; **MAMPs** microbial-associated molecular patterns; **NS** non-suppressor; **Pgt** *Puccinia graminis* f. sp. *tritici*; **PR** pathogenesis-related; **PRRs** pattern recognition receptors; **PTI** pattern-triggered immunity; **R-genes** race-specific genes; **RLCK** receptor-like cytoplasmic kinases; **RLKs** receptor-like kinases; **RNA-Seq** RNA-Sequencing; **SA** salicylic acid; **SAR** systemic acquired resistance; **SNPs** single nucleotide polymorphisms.

INTRODUCTION

Because of their substantial destructive effects on various important crops, rust fungi have been widely studied, and their development in the host plant is well described (Voegele et al., 2009). Under optimal temperature and humidity conditions, urediniospores that land on the leaf surface germinate and form a germ tube. Upon reaching a stoma, the germ tube differentiates into an appressorium from which a penetration hypha is formed, which enters the leaf through the stomatal opening. Within the stomatal cavity, a substomatal vesicle is formed from which an infection hypha emerges. A haustorial mother cell is differentiated upon contact with a mesophyll cell, in which a haustorium is developed. The latter is not intracellular

but is enveloped by the extrahaustorial membrane, derived from the invaginated host plasma membrane. Haustoria play essential roles in nutrient uptake as well as in the synthesis and delivery of effectors, which modulate host metabolism and immune system (Voegele and Mendgen, 2011; Rafiqi et al., 2012). To counteract the pathogen, plants evolved two immune systems: i) Pattern-triggered immunity (PTI) initiated by the recognition of microbial-associated molecular patterns (MAMPs) using pattern recognition receptors (PRRs) and ii) effector-triggered immunity (ETI) initiated by recognition of effectors using resistance (R)-proteins (Dodds and Rathjen, 2010).

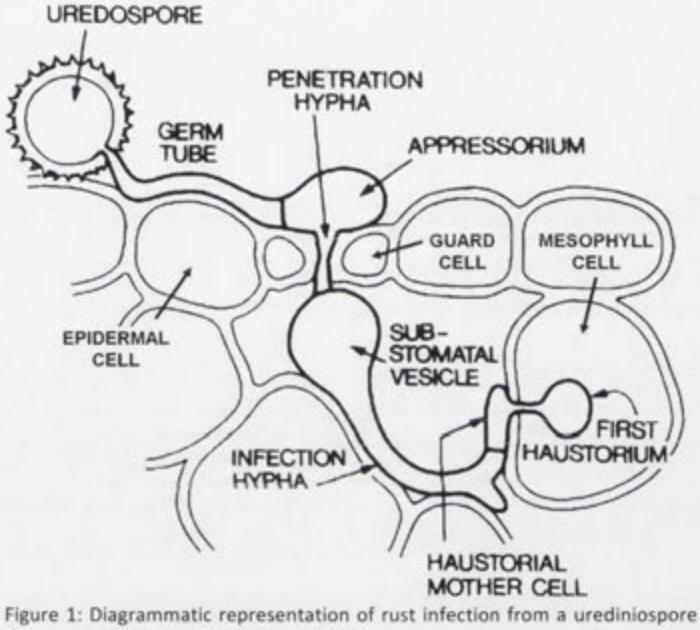


Figure 1: Diagrammatic representation of rust infection from a urediniospore (or uredospore) on the leaf surface. Source: modified from Mendgen (1997).

The utilisation of R-proteins conferring resistance against wheat stem rust, in a race-specific manner for most, is currently the most effective approach for controlling one of the most destructive plant diseases. Global efforts are being made in order to discover new resistance genes and to develop molecular markers linked to these genes, which greatly helps their introgression into agronomically important cultivars. Although, the mechanisms of ETI have been an intense field of study in recent years, many questions remain about the molecular mechanism of resistance that these R-proteins trigger, notably in wheat where little has been done. With an ever-evolving pathogen, which is capable of overcoming any particular resistance genes in a matter of years, characterizing the resistance mechanism may be the key for developing efficient and durable strategies to combat stem rust.

In ETI, the defence response is typically characterized by the hypersensitive response (HR) at the site of infection, which is associated with the death of host cells. Consequently, biotrophic pathogens, such as rusts, are prevented from colonizing the plant further. The resistance response is usually correlated with changes in the expression of host genes. Although genes that are differentially expressed (DE) do not represent the full resistance spectrum, characterizing them is the first step in order to reveal the molecular basis of the resistance mechanism. RNA sequencing (RNA-Seq) has become the tool of choice for gene expression analysis, due in part to the possibility to accurately quantify and qualify whole transcriptomes at once (Jain, 2011). However, few whole-transcriptome analyses have been undertaken so far, which characterize the interaction between wheat and its pathogens in general, and none for wheat and stem rust.

In order to gain some insights into the mechanism of resistance conferred by the 7AL resistance locus, the susceptible line, Columbus (Col), and the resistant lines, Columbus-NS765 (Col-NS765) and Columbus-NS766 (Col-NS766), were inoculated with stem rust (*Puccinia graminis* f. sp. *tritici*, abbreviated as *Pgt*) race #313, and examined over several days post-inoculation (DPI). In a first phase, the development of stem rust *in planta* was observed using fluorescein isothiocyanate (FITC)-labeled wheat germ agglutinin (WGA). WGA binds specifically to N-acetyl-glucosamine (i.e., chitin) and has been used for staining fungal material (Allen et al., 1973; Meyberg, 1988). Coupled with the fluorophore FITC, it allows observing fungal structures easily. In addition, since phenotypic observations showed that the resistance response included cell death (necrosis), cell death was observed using Trypan blue staining, which specifically colours dead cells in blue (Strober, 2001). To go further into the characterization of the resistance mechanism, the transcriptome profiles of Col, Col-NS765 and Col-NS766, were examined using RNA-Seq. Based on the histological observations, three time points were investigated by RNA-Seq: 0, 2 and 5 DPI. 0 DPI was used as the reference time point, while 2 and 5 DPI were used in order to detect the early and late resistance responses, respectively. Genes that were DE between Col and both Col-NS765 and Col-NS766 were considered to be putatively involved in the resistance. For the sake of readability, the term Col-NS is used to refer to Col-NS765 and Col-NS766 collectively.

RESULTS

I. Histological observations

I.1. WGA-FITC

At 0 DPI, urediniospores were seen on the leaf surface (Figure 2.0). At 1 DPI, most spores germinated (Figure 2.1), but very few already developed an appressorium. No difference between genotypes was observed at these time points. At 2 DPI, appressoria were formed over stomata (Figure 2.2), and the pathogen started penetrating the leaf. Interestingly, some guard cells on the resistant lines were autofluorescent under UV and emitted a blue light (Figure 2.2R2). This phenomenon was always associated with an infection site (i.e. the location where appressoria formed, typically at the stoma), and penetration hyphae were often present. Although also seen on Col, autofluorescence was less frequent. At 5 DPI, the rust was well established inside the leaf, in Col, and many hyphae were observed, forming a small network (Figure 2.5S). Haustoria were also visible. For the resistant lines, few hyphal networks were seen, but most fungi were still at the penetration stage. Epidermal cells adjacent to these infection sites were often autofluorescent under UV (Figure 2.5R). Some mesophyll cells were also autofluorescent. At 7 DPI, hyphal networks spread further in Col (Figure 2.7S), whereas observations in the resistant lines were similar than at 5 DPI (Figure 2.7R). Finally, at 9 DPI, hyphal networks were very large in Col, and uredia were sometimes formed (Figure 2.9S). Autofluorescence was emitted all over the networks. For the resistant lines, observations at 9 DPI were similar to those at 5 and 7 DPI. The hyphal networks that developed in the resistant lines were usually smaller than in Col (Figure 3B).

Observations in the susceptible Canthatch-K (CTH-K), were comparable to those for Col, and in the resistant lines, Canthatch-NS1 (CTH-NS1) and Canthatch-NS2 (CTH-NS2), were comparable to Col-NS765 and Col-NS766. At 9 DPI, more and bigger hyphal networks were present on the susceptible lines than on the resistant ones (Figure 3). Also, more rusts were stopped at the penetration stage (i.e. did not developed hyphal networks) in lines with Col background than in lines with CTH background, reflecting the slightly higher susceptibility in the CTH background (Figure 3A).

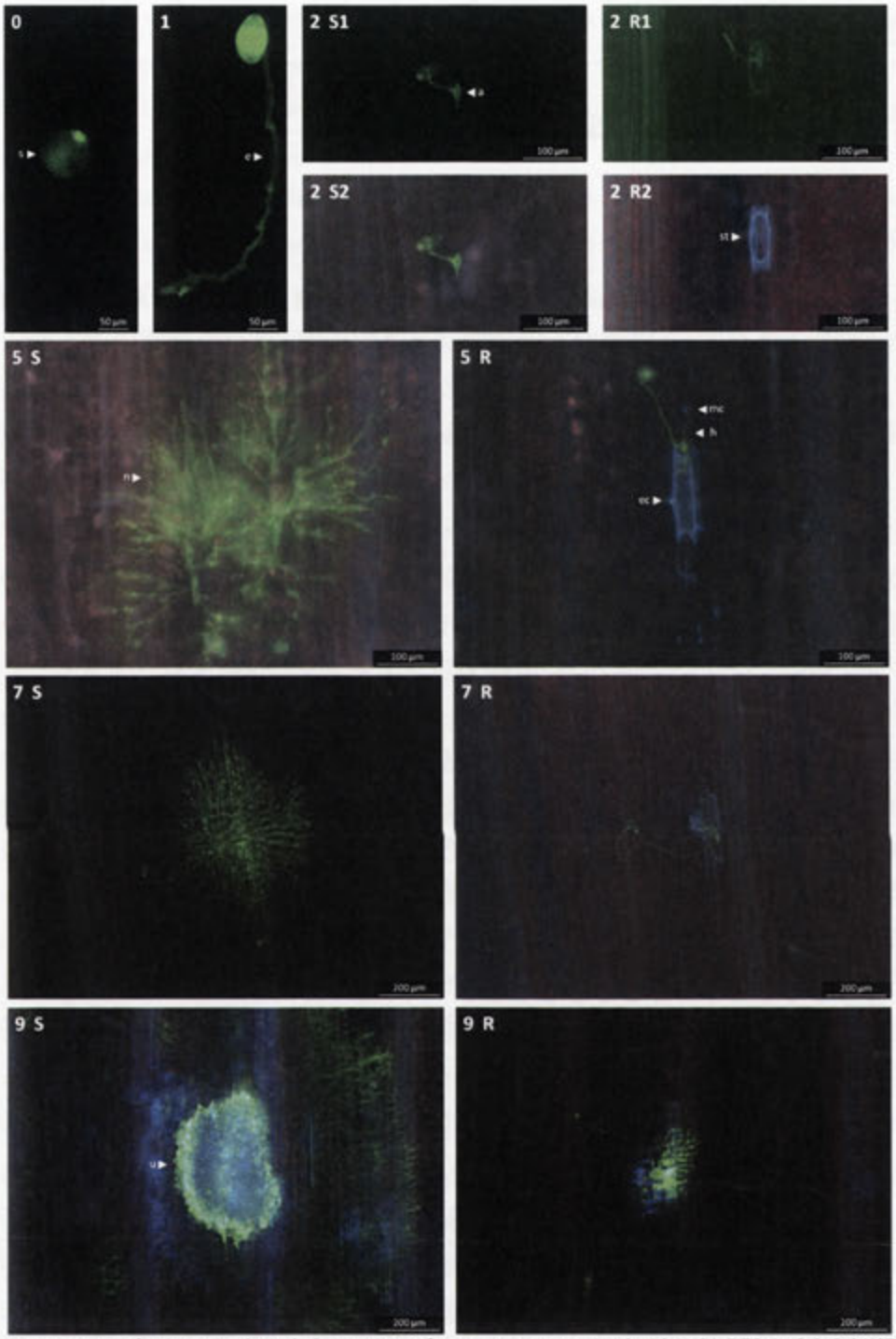


Figure 2: Rust infection. Histological observations of rust structures, stained with WGA-FITC, in Col, Col-NS765 and Col-NS766 at different DPI. Time points are indicated by the corresponding numbers (DPI), and S and R represent the susceptible and resistant typical responses, respectively. Arrows indicate various rust and plant structures: spore (s), elongation tube (e), appressorium (a), hypha (h), hyphal network (n), uredium (u), stoma (st), epidermal cell (ec), mesophyll cell (mc). Pictures 0, 1, 2S1 and 2R1 have been taken under blue light (excitation 470-440 nm, emission 525-550 nm), and the others under UV (excitation 365 nm, emission ≥ 420). 2S1/2S2 and 2R1/2R2 show the same field of view.

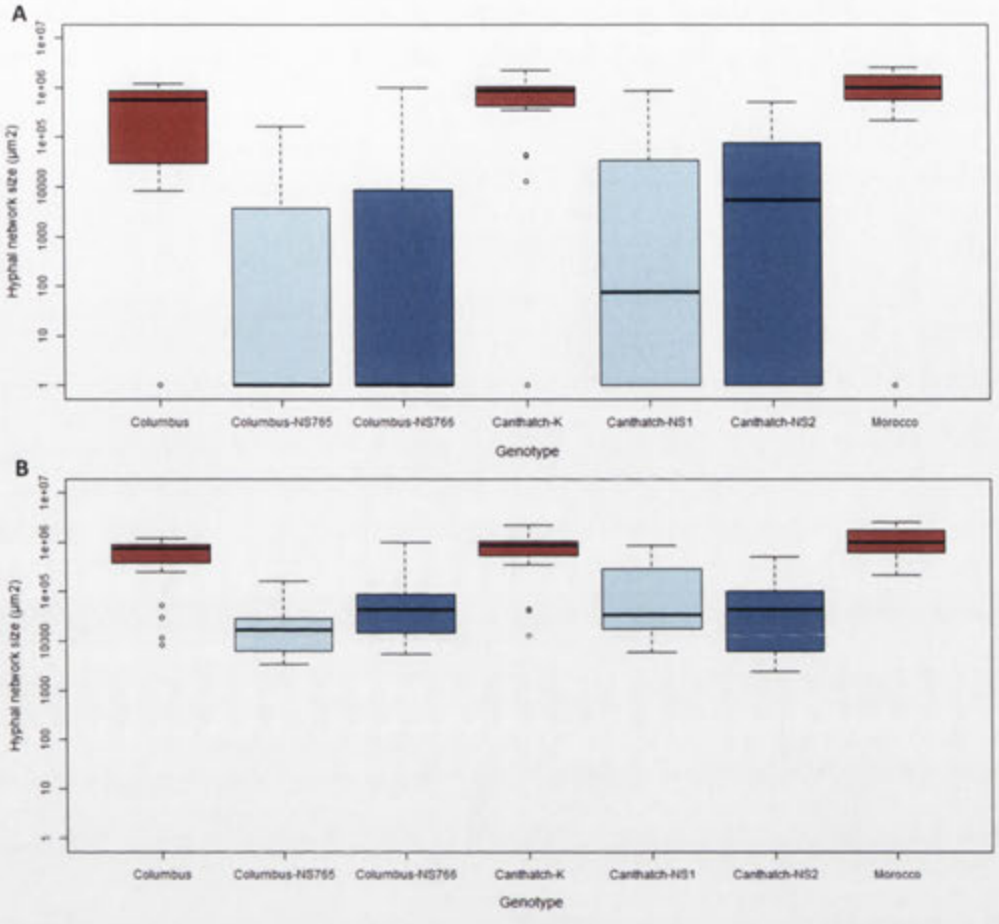


Figure 3: Hyphal network size. Boxplot of the hyphal network size at 9 DPI, in each genotype, with (A) or without (B) counting fungi that were stopped at the penetration stage. The horizontal dark line represents the median and is present in A) at the bottom of the boxes for Col-NS765 and Col-NS766, due to the high number of infection sites that did not develop hyphal networks. Morocco is a complete susceptible line used as a positive control. More and bigger hyphal networks were present on the susceptible genotypes (red) than in the resistant genotypes (blue). Sizes were measured for 30 infection sites on two leaves.

I.2. Trypan Blue

Conveniently, rust spores, elongation tubes and appressoria were often visible after trypan blue (TB) staining (Figure 4). Spores were observed on the leaf surface at 0 DPI and they had germinated at 1 DPI. Some already developed an appressorium over stomata. The first sign of dead plant cells was observed at 2 DPI (Figure 4.2R). Guard cells were sometimes TB stained in the resistant lines but not or partially in Col. This colouration seemed always associated with the presence of spore/appressorium. At 5 DPI, epidermal cells adjacent to infection sites were stained, especially in the resistant genotypes (Figure 4.5R). At 9 DPI, uredia surrounded by many TB stained epidermal cells were observed in Col (Figure 4.9S), but were

not seen in the resistant lines (Figure 4.9 R). The coloration, indicating cell death, at this stage in Col was probably due to the extensive damage caused by the rust development and the formation of uredinia that broke through the plant epidermis.

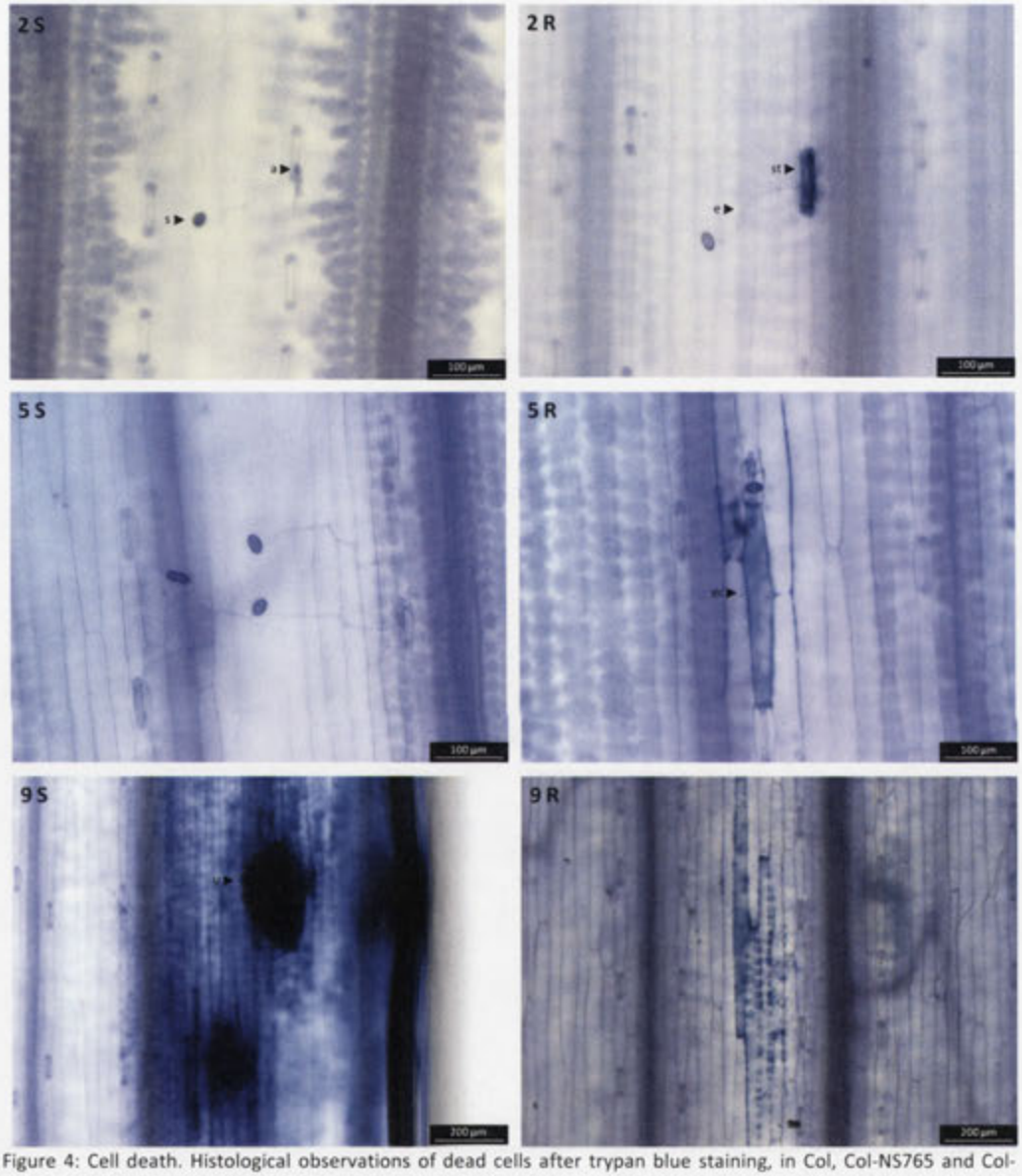


Figure 4: Cell death. Histological observations of dead cells after trypan blue staining, in Col, Col-NS765 and Col-NS766 at different DPI. Time points are indicated by the corresponding numbers (DPI), and S and R represent the susceptible and resistant typical responses, respectively. Arrows indicate various rust and plant structures: spore (s), elongation tube (e), appressorium (a), uredium (u), stoma (st), epidermal cell (ec). Hyphal networks were sometimes visible, but cannot be compared to observations with WGA-FITC staining as Trypan Blue staining is not specific to rust structures.

II. Transcriptome analysis

II.1. Alignment

Total RNA was isolated from samples consisting of the first-leaf of two seedlings at 0, 2 and 5 DPI. Sequencing of mRNA yielded 36 millions reads on average per sample, and a total of 982 millions reads, corresponding to 27 samples (3 genotypes x 3 time points x 3 biological replicates). Reads were aligned against the wheat Unigene database, corresponding to 158k transcripts. Only reads that aligned against unique Unigenes (best hit) were used for differential expression analysis. Alignment yield averaged at 65 % of the total read set. The correlations of read counts for each Unigene, between samples, were calculated (Figure 5). All samples were positively correlated between each other (average 91 % correlation), especially between biological replicates, which were almost perfectly correlated (average 98 % correlation). However one sample, which corresponds to the third replicate of Col-NS765 at 0 DPI, was not as well correlated with any other sample. Because this sample showed a correlation of only 77 % with the corresponding replicates, it was removed from the analysis.

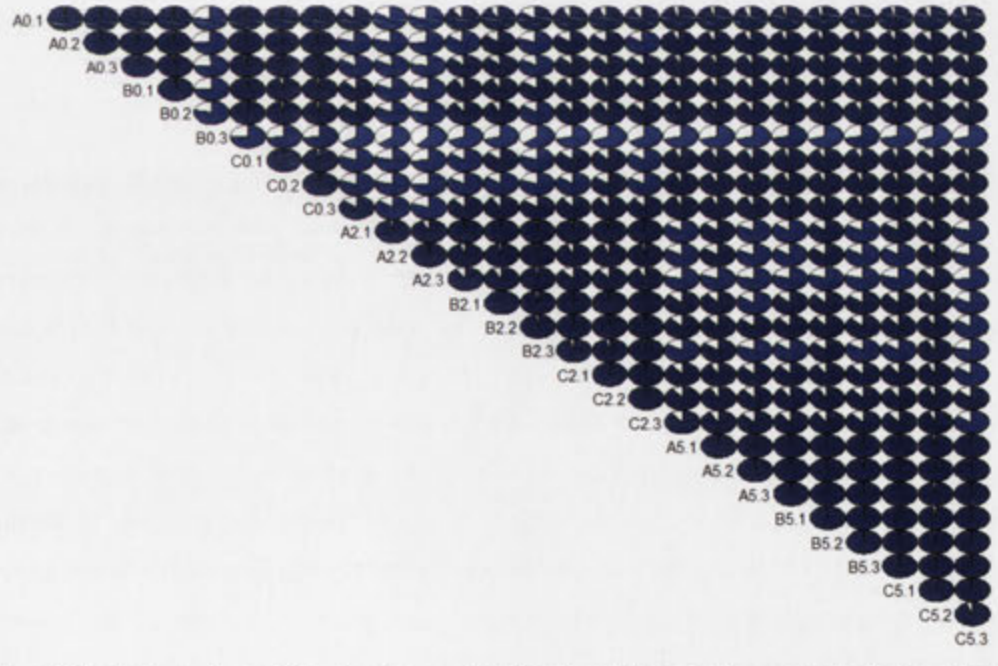


Figure 5: Reads alignment correlation. Correlogram of read counts per Unigene between samples. Each correlation is shown at the intersection of two samples, by a coloured circle. The filled area and colour intensity are proportional to the positive correlation. Letters represent genotypes (A: Col, B: Col-NS765, C: Col-NS766). The first numbers represent time points (0: 0 DPI, 2: 2 DPI, 5: 5 DPI) and the last numbers represent biological replicates.

Unigenes were filtered to those with a minimum of 100 reads across samples, resulting in a total of 67,156 Unigenes expressed across samples. Of these Unigenes, more than half had less than 30 aligned reads for each sample, and around 75 % had less than 100 aligned reads (Figure 6). Less than 30 % of the total reads were found in 95 % of the Unigenes. This implied that most of the sequenced reads were associated with few genes that were highly expressed. These Unigenes corresponded to the "RuBisCO activase", with 500k reads (2.5 %) in average per sample, and various chloroplast and mitochondrial related genes.

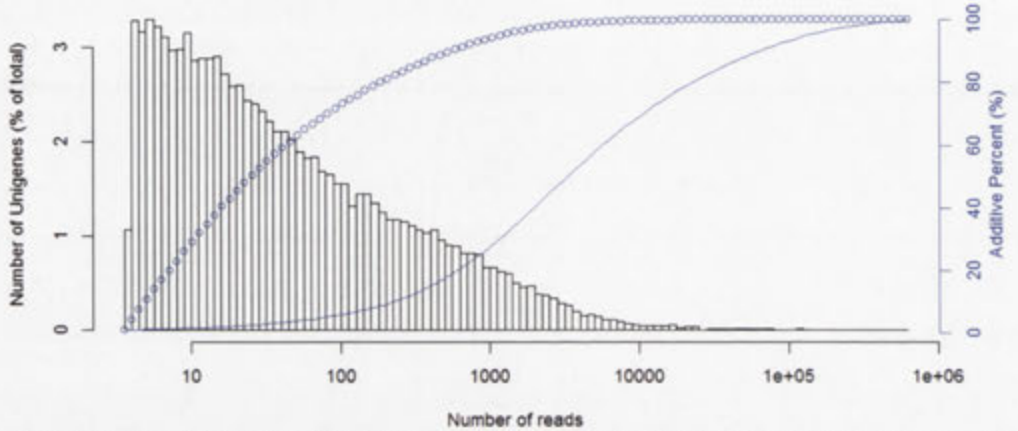


Figure 6: Number of Unigenes relative to the number of reads. Bars represent the percent of Unigene for a particular number of reads, and blue circles represent the additive percent of these Unigene. The blue line represents the additive percent of reads.

II.2. Differential expression analysis

Read counts were first normalised with DESeq (Anders and Huber, 2010). A clustering analysis between samples was performed in order to check the relationship between them (Figure 7). Three main clusters were clearly distinct, each representing a different time point. Genotypes were also grouped into 3 sub-clusters at 2 and 5 DPI, whereas this was not the case at 0 DPI. In addition, Col-NS765 and Col-NS766 were closer to each other, at 2 DPI, than to Col. DE genes were then detected with DESeq. 18 pairwise comparisons were performed, and grouped into two categories: i) genotype comparisons at each time point and ii) time comparisons for each genotype. Genes were considered DE at a false discovery rate (FDR) ≤ 0.05 and an absolute value of fold change (FC) ≥ 2. On average, 8,130 (12.1 %) and 470 (0.7 %) genes were DE between time points and between genotypes, respectively. Once DE genes were determined for a particular condition (e.g. 0 vs 2 DPI in Col), these same genes were re-evaluated in similar conditions (e.g. 0 vs 2 DPI in Col-NS766), using new thresholds (FDR ≤ 0.075, |FC| ≥ 1.8). This adjustment was performed to detect genes that were found DE in one

condition but not in another, although the changes in their expression were similar. For instance, Ta.34224 was found DE between 0 and 2 DPI in Col-NS765 and Col-NS766 (FDR ≤ 0.05, FC = 2.92 and 2.07 respectively), but not in Col (FDR ≤ 0.05, FC = 1.97), although their FC were very similar.

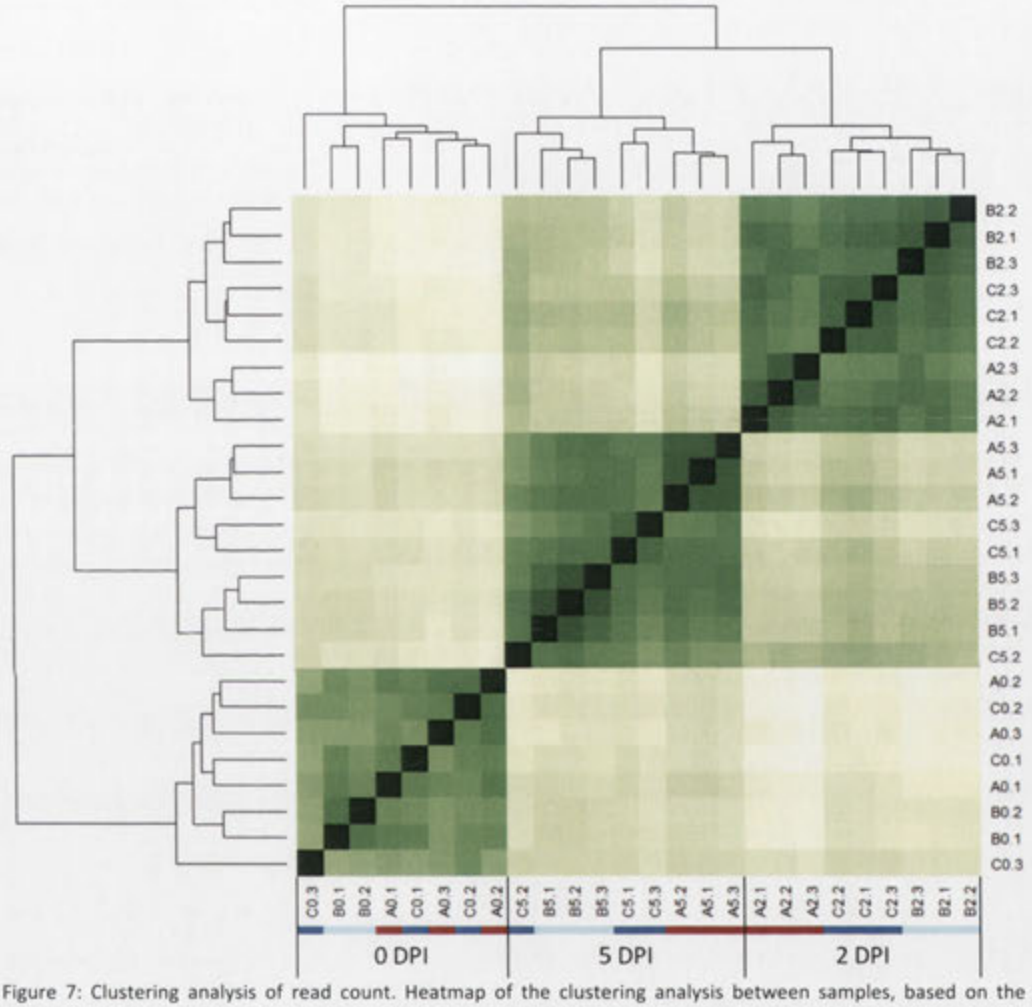


Figure 7: Clustering analysis of read count. Heatmap of the clustering analysis between samples, based on the normalised reads count for each Unigene. Letters and coloured horizontal bars represent genotypes (A: Col, red; B: Col-NS765, pale blue; C: Col-NS766, dark blue). The first numbers represent time points (0: 0 DPI, 2: 2 DPI, 5: 5 DPI) and the last numbers represent biological replicates.

Conditions for calling DE genes were not made too stringent (e.g. FDR ≤ 0.01, FC ≥ 3) in order to maximise the number of candidate genes. Indeed, this experiment was based on the transcriptome profile of whole leaves, but as shown by the histological observations, the transcriptomic changes due to rust infection, were expected to be localised to the few cells flanking infection sites, at least at 2 DPI. The gene expression levels in these cells may have

been diluted by the ones in the other cells, making their detection more difficult. The adjustment of DE genes with secondary thresholds was also done in this perspective.

II.2.1. Validation

In order to make sure that gene expression levels were comparable between samples, read counts for 12 wheat reference genes, selected from a study by Paolacci et al. (2009) were evaluated (Table 1). None of these genes were found DE between genotypes, and only two (Ta.16204 and Ta.44405) were down-regulated at 5 DPI, compared to 0 DPI, in all genotypes (Figure Sup. 1). These results demonstrated a lack of underlying bias in the data because no differences were found in expression of these reference genes between genotypes and expression for most genes was stable through time. Thus the analysis was considered valid.

Unigene	GenBank	Annotation
Ta.46201	JV864494	Cell division control prot., AAA-superfamily of ATPases
Ta.2291	JV906091	ADP-ribosylation factor
Ta.2776	JP833794	RNase L inhibitor-like protein
Ta.35284	JV864644	Protein transport protein SEC23-like
Ta.22845	JV883866	Proteasome subunits
Ta.50503	JV910606	Ubiquitin
Ta.25534	DQ435668	α-tubulin
Ta.16204	JV884210	GAPDH
Ta.54825	HP620998	Actin
Ta.38797	DR738338	Histone
Ta.44405	U76895	β-tubulin
Ta.27771	CK168306	Ribosomal protein

Table 1: Wheat reference genes.

II.3. Genes of interest

More than 15,000 and 11,000 genes were found DE in at least one genotype, at 2 and 5 DPI, respectively, compared to 0 DPI (Figure 8 Top). Around 40 % of them were regulated the same way in all genotypes and several thousand were DE in only one genotype. More genes were DE in Col at 2 and 5 DPI, compared to 0 DPI, than in Col-NS. These results indicated that several thousand genes were potentially DE between genotypes. However, results from the comparisons between genotypes at each time point, showed that it was not the case (Figure 8 Bottom). A total of 1,500 genes were actually DE between at least two genotypes, at any time point. The indirect detection of DE genes between genotypes, using the differences in their regulation through time for each genotype, was therefore not appropriate. This may be due to the high number of genes that were DE through time, enhancing the chance of false negatives,

particularly for such a low stringency of differential expression parameters. For this reason, only genes that were DE between genotypes were followed up. Across time points, Col had always more DE genes with Col-NS765 (1128 in total) than with Col-NS766 (610 in total). A total of 687 genes were DE between Col-NS765 and Col-NS766.

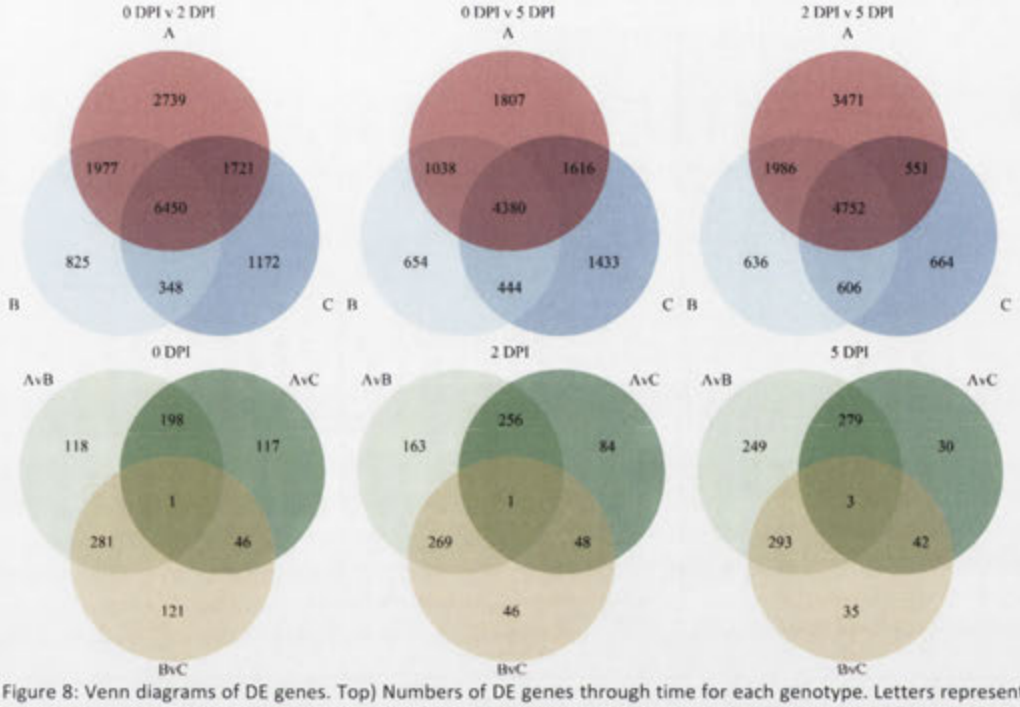


Figure 8: Venn diagrams of DE genes. Top) Numbers of DE genes through time for each genotype. Letters represent genotypes (A: Col, red; B: Col-NS765, pale blue; C: Col-NS766, dark blue). Bottom) Numbers of DE genes between genotypes at each time point (Col vs Col-NS765: pale green; Col vs Col-NS766: dark green; Col-NS765 vs Col-NS766: brown). The genes of greatest interest are those that were DE between the susceptible Col, and both resistant Col-NS765 and Col-NS766, that is, at the intersection of "AvB" and "AvC".

Genes that were DE between Col and both Col-NS lines, at any time point, were selected as genes of interest. These genes were potentially involved in the resistance response. A total of 353 genes were selected, of which 198, 256 and 279 were DE between Col and both Col-NS lines at 0, 2 and 5 DPI, respectively. At any time point, there were always more genes that were more expressed in Col-NS vs. Col, than genes that were less expressed (Figure 9). Never genes were DE one way at one time point, and the other way at a different time (e.g. more expressed at 2 DPI and less expressed at 5 DPI). More than half of the genes of interest were either more, or less expressed in Col-NS vs. Col by a fold change of minimum eight (Figure 9A). Fold changes were usually similar between Col and either Col-NS765 or Col-NS766. Nearly half of the genes of interest were common to all three time points, and more than 80 % of the genes DE at 0 DPI were also DE at 2 and 5 DPI (Figure 9B). Among the genes in common,

70 % of them were never DE through time in any genotype. Moreover, 65 % and 45 % of the common genes were never expressed in Col and Col-NS, respectively.

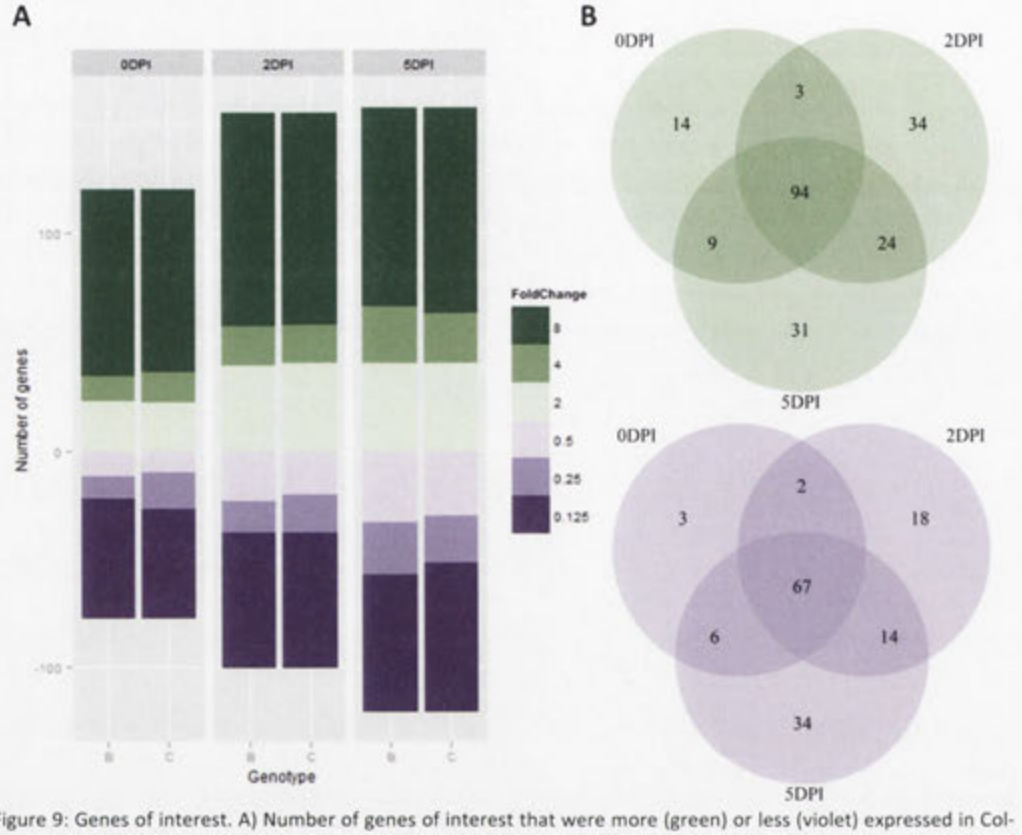


Figure 9: Genes of interest. A) Number of genes of interest that were more (green) or less (violet) expressed in Col-NS765 and Col-NS766, compared to Col, at each time point, with fold change represented by intensity of colour. Letters represent the genotype compared to Col (B: Col-NS765; C: Col-NS766). B) Venn diagrams of genes of interest that are more (green) or less (violet) expressed in Col-NS, compared to Col, at each time point.

Unigene sequences were mapped to chromosomes, by comparing the Unigene reference sequences with the wheat genome survey sequencing (GSS) database. More than 99% of the Unigenes used for the DE analysis (67,156 Unigenes) were assigned to chromosomes. The Unigenes were found to be uniformly distributed across the genome (Figure 10A). However, among the genes of interest that were DE at 0 DPI, many were found located on chromosomes 6AL, 6AS and 7AL, which carried 41 (20.7 %), 18 (9.1 %) and 16 (8.1 %) genes, respectively (Figure 10B). 19 genes (9.6 %) were also found on 5BL, but as there were more Unigenes on this chromosome than on the others (Figure 10A), the percent of Unigenes per chromosome that were DE was actually not as high as for the others listed above.

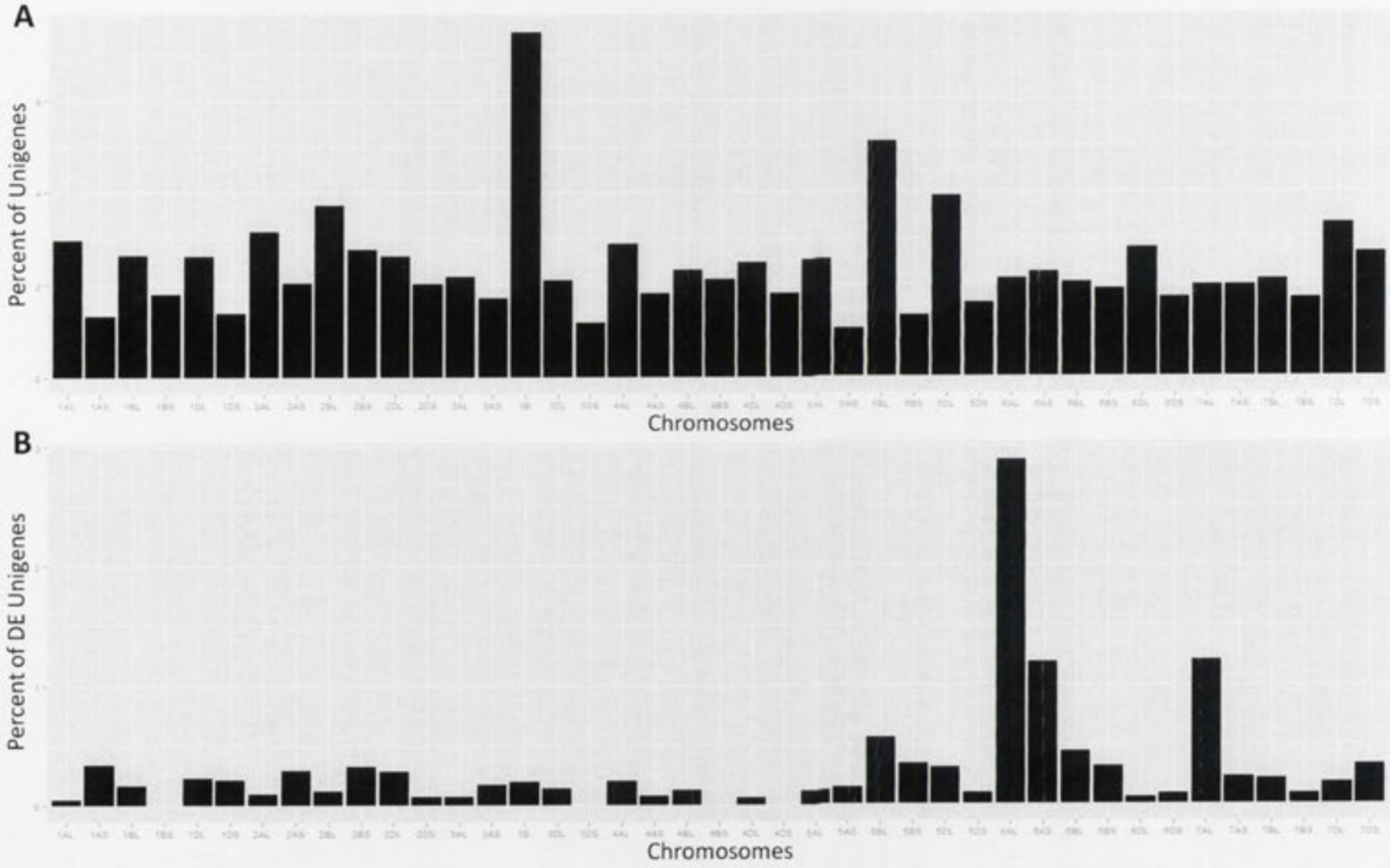


Figure 10: Chromosomal location of Unigenes. A) Percent of total Unigenes with reads aligned, mapped to each chromosome. Note: chromosome 3B includes both 3BL and 3BS. B) Percent of Unigenes mapped to each chromosome that were DE between Col and Col-NS, at 0 DPI.

The 353 genes of interest were classified into various categories (Table 2), depending on the time points at which they were found DE between Col and Col-NS, and if they were DE through time. Genes could belong to more than one category. In some cases, genes did not meet the criteria used for a particular category because they failed to meet the statistical criteria of the DESeq test, but could obviously be included in this category. For instance, the genes Ta.154160 and Ta.160425 were not at first classified in the "origin" category because the FDR at 0 DPI of the comparison between Col and Col-NS was too high, although these genes were never expressed in Col but always in Col-NS (Figure 11). Due to this adjustment, genes could sometimes be classified as constitutively expressed (i.e. constitutive category) and inducible (e.g. "both" category), which is contradictory. In this case, genes were classified as "constitutive".

Almost half of the genes were found to be constitutively expressed, that is, always DE between Col and Col-NS, and never DE through time in any genotype (Table 2). 63 and 126 genes were categorised as putatively involved in the early and late resistance response, respectively, and 46 genes in both resistance responses. Genes were annotated using the rice, Arabidopsis and NCBI Refseq databases. Among the genes of interest, 113 (32 %) were annotated (Table 3), 200 (57 %) were poorly or not annotated, and 40 (11 %) were repetitive or transposable elements. All gene classifications and annotations can be viewed in File Sup. 1.

Category	Definition	Criteria	More expressed in Col-NS		Less expressed in Col-NS	
			#	%	#	%
Origin	Genes putatively at the origin of the resistance/susceptible response (e.g. R gene).	DE at 0 DPI between Col and Col-NS, and located on chromosome 7.	19	5%	10	3%
Constitutive	Genes constitutively expressed.	Always DE between Col and Col-NS, and never DE through time in any genotype	92	26%	61	17%
Early	Genes putatively involved in the early resistance response.	DE at 2 DPI between Col and Col-NS, but not at 0 DPI unless DE at 2 DPI, compared to 0 DPI.	38	11%	25	7%
Late	Genes putatively involved in the late resistance response.	DE at 5 DPI between Col and Col-NS, but not at 0 DPI unless DE at 5 DPI, compared to 0 DPI.	46	13%	34	10%
Both	Genes putatively involved in both early and late resistance responses.	DE at 2 and 5 DPI between Col and Col-NS, but not at 0 DPI unless DE at 2 and 5 DPI, compared to 0 DPI.	7	2%	9	3%
Pathogen	Genes putatively from the pathogen.	Not expressed at 0 DPI in any genotype, and up-regulated in Col at 5 DPI, compared to 0 DPI.	-	-	4	1%
Zero	-	DE at 0 DPI between Col and Col-NS, but not at 2 and 5 DPI.	13	4%	3	1%
Miscellaneous	-	Genes that do not fit the above conditions.	8	2%	6	2%

Table 2: Classification of genes of interest. Criteria used for categorizing the genes found DE between Col and both Col-NS765 and Col-NS766, and the number and percentage of the 353 genes of interest in each category. Note: these criteria are not restrictive and in some cases, genes could be classified in a particular category even though they did not perfectly meet all criteria. Graphs of gene expression are presented for genes in most categories in the discussion.

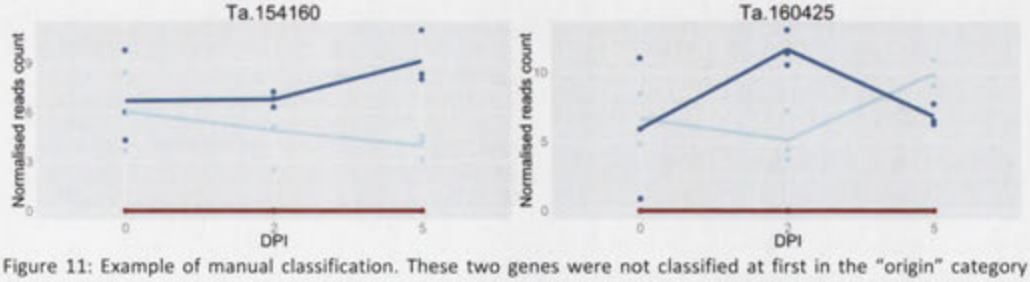


Figure 11: Example of manual classification. These two genes were not classified at first in the "origin" category because they were not found statistically DE at 0 DPI although they were never expressed in Col but always, at a low level, in Col-NS. For each Unigene, the DESeq-normalised reads count is reported for each sample (dots) at the different time points. Lines represent the biological replicates average. Colours represent genotypes: Col (red), Col-NS765 (pale blue), Col-NS766 (dark blue).

Unigene	Col_0DPI	Col-NS765_0DPI	Col-NS766_0DPI	Col_2DPI	Col-NS765_2DPI	Col-NS766_2DPI	Col_5DPI	Col-NS765_5DPI	Col-NS766_5DPI	Annotation	Chromosome	Col-NS/Col
Origin												
Ta.75386	0.9	3.3	5.2	1.7	9.2	8.6	0.7	5.1	5.6	CRK	7AL	U
Ta.90783	3.4	6.6	6.1	4.3	10.5	9.1	2.0	10.6	9.8	CRK	7AL	U
Ta.90847	0.0	0.5	0.8	0.0	1.6	1.3	0.0	1.7	1.5	CRK	7DL	U
Ta.95070	0.0	0.6	0.7	0.1	1.3	1.1	0.0	0.9	0.9	CRK	7AL	U
Ta.104804	0.0	14.6	14.8	0.0	24.8	23.5	0.1	20.0	21.5	CRK	7AL	U
Ta.109657	0.1	2.8	3.7	0.3	7.8	7.0	0.1	4.1	4.3	CRK	7BL	U
Ta.6285	7.3	3.7	3.6	5.9	3.7	3.8	8.3	4.4	3.9	Ribonuclease	7AL	D
Ta.42512	1.6	0.0	0.2	35.8	1.9	1.2	13.7	0.5	0.4	CYP	7DL	D
Ta.50162	5.0	2.2	1.4	11.7	5.4	5.1	9.4	3.3	3.4	Inorganic H+ pyrophospha-tase	7AL	D
Ta.77864	2.0	0.1	0.2	3.0	0.9	0.3	1.4	0.4	0.3	Coatomer subunit beta	7AL	D
Ta.140170	1.4	0.1	0.2	2.8	0.4	0.3	1.5	0.2	0.2	F-box/LRR	7DL	D
Constitutive												
Ta.4126	2.2	5.3	3.9	3.6	7.5	6.4	1.8	5.0	4.7	XH/XS domain	3AL	U
Ta.10950	0.8	2.2	1.4	0.8	3.2	2.2	0.8	2.8	2.0	Apomixis-related	7BS	U
Ta.11030	1.9	3.6	3.8	3.4	5.5	6.0	1.7	3.6	3.6	ATPase	6AL	U
Ta.43720	1.8	5.1	4.4	2.5	6.0	5.5	2.3	6.6	5.7	SNF2 domain	5DL	U
Ta.79662	0.0	1.2	1.4	0.0	1.5	1.8	0.0	1.3	1.6	MAP kinase	2AS	U
Ta.83884	0.0	2.3	1.8	0.0	3.7	3.6	0.0	2.6	2.9	Transportin 1-like	2BL	U
Ta.85159	0.0	2.4	1.5	0.0	3.7	2.2	0.0	1.9	1.2	RNA-binding	7DS	U
Ta.95070	0.0	0.6	0.7	0.1	1.3	1.1	0.0	0.9	0.9	CRK	7AL	U
Ta.99566	0.1	4.0	4.1	0.2	3.2	3.8	0.1	3.0	2.8	Methyltransferase	3B	U
Ta.104804	0.0	14.6	14.8	0.0	24.8	23.5	0.1	20.0	21.5	CRK	7AL	U
Ta.111478	0.4	1.0	1.1	0.6	2.7	2.3	0.4	1.4	1.4	CRK	7DS	U
Ta.115331	0.0	0.2	0.3	0.0	0.5	0.6	0.0	0.5	0.3	MaoC like domain	6AL	U
Ta.155345	0.1	2.7	2.5	0.1	3.1	3.2	0.0	3.9	3.3	CRK	5DL	U
Ta.6285	7.3	3.7	3.6	5.9	3.7	3.8	8.3	4.4	3.9	Ribonuclease	7AL	D
Ta.12047	2.3	0.3	0.1	3.0	0.1	0.2	3.2	0.1	0.1	Disease resistance protein	4AL	D
Ta.18874	14.0	1.3	0.5	19.4	0.8	0.4	11.4	0.6	0.3	NBS-LRR disease resistance	4AL	D
Ta.39158	23.0	0.0	0.0	27.6	0.0	0.0	18.3	0.1	0.0	Disease resistance protein	7DS	D
Ta.72561	6.5	0.5	0.5	6.8	0.4	0.7	7.9	0.5	0.7	Myosin-like	6AL	D
Ta.81001	4.8	0.4	0.6	3.5	0.3	0.8	6.0	0.4	0.5	Ulp1 protease	2DL	D
Ta.91588	2.4	0.0	0.0	4.0	0.0	0.0	2.5	0.0	0.0	FAR1-RELATED SEQUENCE	6AL	D
Ta.93375	1.6	0.2	0.0	2.3	0.0	0.1	1.5	0.0	0.0	NBS-LRR disease resistance	5BL	D
Ta.93435	1.8	0.3	0.4	2.1	0.4	0.5	4.0	0.4	0.4	COBRA-like	6BL	D
Ta.94685	6.0	0.0	0.0	5.8	0.0	0.0	8.0	0.0	0.0	Nuclear-pore anchor	6AS	D
Ta.103356	1.3	0.1	0.0	1.5	0.0	0.0	1.8	0.1	0.0	Ankyrin repeats motif	5BL	D
Ta.108580	2.9	0.0	0.0	2.5	0.0	0.0	2.8	0.0	0.0	Ankyrin repeats motif	5BL	D
Ta.140170	1.4	0.1	0.2	2.8	0.4	0.3	1.5	0.2	0.2	F-box/LRR	7DL	D
Ta.169085	3.9	0.9	0.9	6.6	1.8	1.6	6.1	1.5	1.5	Transcription factor	6AS	D
Ta.174313	12.7	0.0	0.0	17.5	0.0	0.2	10.6	0.0	0.0	NBS-LRR disease resistance	5BL	D
Early												
Ta.1164	3.0	4.6	4.4	1.4	2.6	2.8	2.8	3.9	4.2	Serine/threonine-protein kinase	1BL	U
Ta.7579	0.7	5.6	4.5	0.4	6.8	6.2	0.5	9.8	8.5	Transporter	3B	U
Ta.25495	0.3	1.1	0.6	0.4	1.8	1.4	0.1	0.5	0.3	Deoxyuridine 5-triphosphate nucleotidohydrolase	2AL	U
Ta.45840	3.4	6.4	5.0	4.3	7.7	9.0	4.4	7.2	7.0	Calmodulin-binding protein	1BL	U
Ta.75637	10.7	27.6	12.5	4.7	13.7	8.8	14.0	35.5	22.1	Auxin-responsive	6BL	U
Ta.92767	18.3	45.3	24.6	8.0	25.6	18.2	10.2	23.9	16.6	Abscisic stress-ripening	3B	U
Ta.107584	13.0	19.5	17.2	2.4	6.8	8.8	25.7	38.0	34.0	Pentatricopeptide containing protein	6AS	U
Ta.10772	1.9	1.1	1.2	1.7	0.4	0.4	1.0	0.4	0.5	nodulin MtN2-like	2BS	D
Ta.13785	17.9	17.0	17.6	31.9	12.6	7.0	2.9	3.7	2.5	Xylanase inhibitor	4AL	D
Ta.36124	2.5	1.3	2.7	12.6	2.4	3.6	5.8	3.0	3.5	ADP-ribosylation factor GTPase-activating protein	6AL	D
Ta.38600	1.2	0.7	0.7	4.7	1.5	1.7	1.7	0.8	1.0	B3 DNA binding domain	5AL	D
Ta.39365	2.6	4.1	3.8	2.4	0.2	0.2	0.2	0.3	0.3	Linalool synthase	2DL	D
Ta.50162	5.0	2.2	1.4	11.7	5.4	5.1	9.4	3.3	3.4	Inorganic H+ pyrophospha-tase	7AL	D
Ta.68844	1.8	0.4	0.3	0.9	0.1	0.1	1.1	0.1	0.2	methionine aminopeptidase	6AL	D
Ta.70282	4.0	2.1	3.3	6.0	2.8	2.9	8.1	5.0	5.4	replication factor A	6AL	D
Ta.73172	7.3	4.8	3.2	14.5	6.9	7.9	14.6	8.3	9.7	NIN Like Protein	5BL	D
Ta.74923	2.3	1.8	2.6	2.9	1.2	1.1	2.7	1.6	2.4	endo-1,3;1,4-beta-D-glucanase-like	7BL	D
Ta.75060	2.7	2.2	1.7	5.9	2.4	2.9	3.4	1.9	2.4	Disease resistance	7DL	D
Ta.105098	1.9	2.8	2.7	1.2	0.1	0.1	0.2	0.1	0.2	Linalool synthase	6DS	D
Late												
Ta.9901	2.0	3.3	3.3	3.9	5.4	5.5	2.2	4.5	4.8	bromodomain domain containing protein	5BL	U

Ta.14904	0.4	1.1	1.6	4.5	5.0	2.8	2.8	12.2	11.0	CYP	2BL	U
Ta.22342	0.3	0.4	0.6	0.5	0.5	0.6	0.5	1.0	1.2	mitochondrial ribosomal protein S3	4BL	U
Ta.33469	20.2	26.2	18.7	2.8	2.9	2.2	2.0	5.9	5.1	jumonji class transcription factor	1DL	U
Ta.42697	0.2	0.3	0.3	0.3	0.4	0.6	0.2	0.4	0.5	NADH dehydrogenase	1BS	U
Ta.43952	0.0	0.2	1.4	0.1	0.6	2.2	0.0	0.5	1.9	microtubule associated protein	2DL	U
Ta.45366	0.3	1.4	1.5	0.4	1.3	1.4	0.6	3.4	2.7	APC4	6AL	U
Ta.45678	0.2	5.4	5.6	0.4	5.2	6.9	0.2	4.7	4.1	E3 ubiquitin-protein ligase	3B	U
Ta.54838	0.1	1.6	0.5	1.4	4.4	3.9	1.5	6.3	5.5	HMGB	5BL	U
Ta.66505	0.8	1.1	0.8	2.3	2.9	2.2	1.1	3.0	2.7	NBS-LRR disease resistance	1BL	U
Ta.71904	0.2	1.2	1.4	0.3	1.1	1.2	0.6	2.9	2.8	APC4	6BL	U
Ta.72300	63.8	83.5	52.3	36.4	23.7	21.9	22.0	46.3	52.1	RLK	6DL	U
Ta.75386	0.9	3.3	5.2	1.7	9.2	8.6	0.7	5.1	5.6	CRK	7AL	U
Ta.77626	0.4	11.3	11.8	0.2	11.4	11.8	0.5	19.4	19.2	MATE efflux family protein	5BL	U
Ta.78049	0.6	0.4	0.9	1.1	1.0	0.8	0.8	2.2	2.3	ABC transporter	6AS	U
Ta.87342	1.8	6.4	3.5	3.6	5.7	5.2	2.0	6.3	5.4	Receptor-like ser-ine/threonine protein kinase	6DL	U
Ta.87731	0.3	0.9	0.7	0.6	1.3	1.0	0.6	2.8	2.3	Disease resistance protein	1DL	U
Ta.87792	0.3	0.7	0.3	0.7	1.6	0.7	0.9	2.3	2.3	RLK	2BL	U
Ta.88423	0.1	0.7	0.6	0.5	0.9	0.4	0.5	2.5	2.1	NBS-LRR disease resistance	1AL	U
Ta.89279	3.2	9.0	5.2	4.9	9.1	5.9	2.9	8.1	7.9	serine/threonine-protein kinase receptor	6AL	U
Ta.90050	0.0	0.1	0.4	0.8	0.9	0.9	0.8	2.6	2.8	UDP-glycosyltransferase	2BS	U
Ta.90783	3.4	6.6	6.1	4.3	10.5	9.1	2.0	10.6	9.8	CRK	7AL	U
Ta.101645	0.0	0.3	0.3	0.2	1.3	0.7	0.2	1.5	1.0	CRK	3B	U
Ta.102188	1.0	1.3	0.8	7.4	7.7	4.9	1.1	2.7	2.6	CRK	1AL	U
Ta.103989	1.1	2.1	2.1	2.3	3.8	3.5	1.3	2.6	2.9	CRK	5DL	U
Ta.180798	0.4	1.0	1.4	1.6	3.2	2.9	1.1	2.5	2.8	coatomer subunit beta	7BL	U
Ta.11076	1.2	0.6	1.6	0.7	1.6	1.2	4.3	1.5	1.1	UDP-glycosyltransferase	2BS	D
Ta.12753	0.9	0.7	1.1	0.3	0.2	0.3	1.2	0.4	0.4	acyltransferase	1BL	D
Ta.14588	0.2	0.2	0.3	0.6	0.5	0.4	0.6	0.1	0.1	polygalacturonase inhibitor	2AS	D
Ta.22082	0.0	0.1	0.1	21.5	19.3	16.9	2.1	0.8	0.7	malate synthase	2DL	D
Ta.22245	6.3	3.4	4.5	2.9	1.6	1.8	3.7	1.1	1.7	Ankyrin repeats motif	3DS	D
Ta.31840	2.3	2.7	3.6	1.4	4.4	3.6	2.8	1.1	1.0	basic helix-loop-helix (bHLH) DNA-binding protein	5AL	D
Ta.33173	0.1	0.3	0.1	19.6	13.5	16.8	2.6	0.5	0.4	isocitrate lyase	2BS	D
Ta.35098	18.7	15.5	16.5	0.2	0.3	0.6	2.2	0.7	0.4	9-cis-epoxycarotenoid dioxygenase	5BS	D
Ta.46729	0.7	0.2	0.4	0.3	0.1	0.0	0.5	0.1	0.1	methionine aminopeptidase	6AL	D
Ta.53934	676.8	598.8	738.8	43.7	37.2	42.7	190.7	97.3	93.7	sterol C4-methyl oxidase 1-2	4DS	D
Ta.55220	0.2	0.1	0.2	0.4	0.3	0.3	0.3	0.0	0.0	polygalacturonase inhibitor	2DS	D
Ta.77182	1.6	0.0	0.2	0.8	0.0	0.0	1.5	0.0	0.1	Early-responsive to dehydration stress protein	3DL	D
Ta.102536	2.0	2.0	2.5	11.9	9.7	10.7	7.0	3.8	3.1	alcohol oxidase	1DL	D
Ta.154091	1.1	0.0	0.0	1.4	0.1	0.1	2.8	0.1	0.1	COBRA-like	6AL	D
Ta.166219	0.7	0.1	0.1	0.8	0.1	0.2	0.9	0.2	0.0	Transducin	7DL	D
Ta.195306	9.4	10.5	7.6	9.4	12.7	11.7	35.8	18.5	17.4	CYP	NA	D
Both												
Ta.47554	16.0	4.7	3.6	6.0	1.7	1.9	14.5	5.4	5.3	MADS-box transcription factor	7DS	D
Ta.87395	1.6	0.4	0.5	4.3	0.4	0.4	5.3	1.0	1.1	LRR receptor-like serine/threonine-protein kinase	6AL	D
Ta.94087	5.0	0.0	0.0	2.7	0.0	0.0	7.8	0.0	0.0	GTPase activating protein	6BL	D
Ta.42512	1.6	0.0	0.2	35.8	1.9	1.2	13.7	0.5	0.4	CYP	7DL	D
Ta.100531	29.3	8.3	7.5	10.6	2.2	2.8	23.8	8.6	8.3	MADS-box transcription factor	7DS	D
Pathogen												
Ta.68870	0.0	0.0	0.0	0.5	0.5	0.2	11.6	0.4	0.9	60S ribosomal protein	NA	D
Ta.68873	0.0	0.0	0.0	0.5	0.2	0.3	12.4	0.6	0.7	40S ribosomal protein	NA	D
Ta.68881	0.0	0.0	0.0	0.0	0.1	0.2	1.7	0.1	0.1	40S ribosomal protein	NA	D
Ta.114604	0.0	0.0	0.0	0.1	0.1	0.1	2.4	0.1	0.1	elongation factor Tu	4DL	D

Table 3: Annotated genes of interest. Expression level, in Reads Per Kilobase per Million reads (RPKM), for the Unigenes DE between Col and Col-NS that were annotated, depending on their classification. "U" (Up) and "D" (Down) represent higher or lower expression levels in Col-NS, compared to Col, respectively.

DISCUSSION

I. Microscopy

Through histological observations, the different steps of stem rust development *in planta* were observed, from the landing of spores on the leaf surface to the formation of new spores nine days later. Various rust structures were observed, such as appressoria and haustoria. Differences between resistant and susceptible lines were visible from 2 DPI, and mainly consisted of greater autofluorescence or TB coloration of guard and epidermal cells at sites of penetration by rust in the resistant lines compared to the susceptible. Both autofluorescence and TB coloration are indicative of dead cells (Koga et al., 1988; Strober, 2001). This result suggests that the resistance mechanism involves cell death, predominantly in the epidermal cell layer, and is established as early as 2 DPI. This is in agreement with visual phenotypes recorded at 12 DPI, at which time necrosis was visible around infection sites in resistant lines.

When cell death was observed, it was always associated with an infection site, indicating that the plant only initiates the resistance response once the fungus has developed an appressorium and started penetrating the leaf. The sudden stop of the rust infection due to cell death is usually the sign of HR, which is generally associated with ETI. However, typical HR response involves death of host cells, usually mesophyll cells, after the development of haustorial mother cells and haustoria (Catanzariti et al., 2007). In this study, cell death was mainly associated with guard and epidermal cells, without the development of haustorial mother cells and haustoria. Although the resistance response here is very similar to HR and seems to be part of ETI, it is possible that cell death is due to another mechanism than HR and is part of an amplified response of PTI. Nevertheless, this result suggests that the plant could recognize MAMPs or effectors associated with the hyphae during the penetration stage. Although much less frequent than in Col-NS, cell death was also observed in Col, suggesting that this line also possesses some level of resistance. This is not surprising since Col was not scored, for race #313, as fully susceptible but as moderately susceptible (Figure 12). Interestingly, cell death was also localised to the guard and epidermal cells, indicating that the partial resistance mechanism of Col is similar to that of the resistant lines. The same resistance response may be operating in both Col and Col-NS, but it is somehow enhanced in the latter.

Although not frequent, fungal growth was also observed in the resistant lines, but hyphal networks were usually smaller than in Col (Figure 3). This may be due to the early resistance mechanism during the penetration stage, which did not stop the rust as in most cases, but only delayed its propagation. Another possibility is a late resistance mechanism, which does not involve cell death, such as the systemic acquired resistance (SAR). This late resistance mechanism could have been triggered by the early resistance mechanism and may limit propagation of fungi that evade the early resistance.

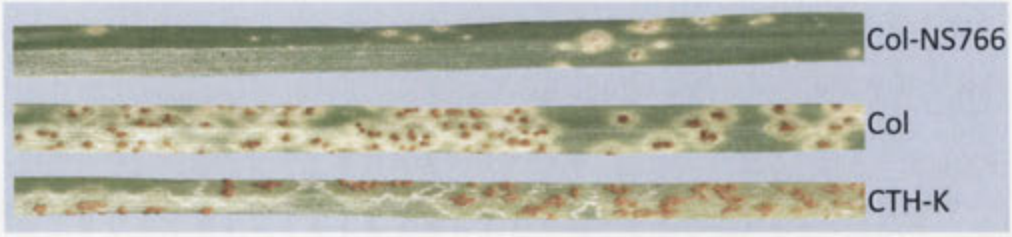


Figure 12: Phenotypic response to stem rust race #313, 15 days post-inoculation on seedlings. Col-NS766 is resistant (IT: ;1), Col is moderately susceptible (IT: 2+3) and CTH-K is nearly fully susceptible (IT: 3+4). See chapter two for more details. IT: Infection Type.

II. RNA-Seq

II.1. Choice of time points

The histological observations at different DPI (0, 1, 2, 5, 7 and 9 DPI) gave some insights into the rust infection and the resistance involved. Moreover, it provided an understanding of the times at which these various processes occurred, which helped in the choice of time points used in the transcriptome analysis. 0, 2 and 5 DPI were considered the optimal time points for studying the resistance mechanism.

- 0 DPI was chosen as the reference time point. This allowed determining the genes that were initially differentially expressed between resistant and susceptible lines, independently of the rust inoculation.
- 2 DPI was chosen for the detection of genes involved in the primary, or early, resistance response. This was the first point at which differences between the responses of the Col and Col-NS genotypes were observed. Although only a minority of infection sites were associated with cell death at this stage, there were a greater number of sites with dead cells in Col-NS than in Col.

5 DPI was chosen for the detection of genes involved in the secondary, or late, resistance response. Hyphal networks were on average smaller in Col-NS than in Col (Figure 3B). Pathogen growth could have been slowed down by the early resistance mechanism at the penetration stage, but also by another mechanism during fungal colonisation of the plant. Transcriptome profiling of plants with developed hyphal network could help in determining this.

II.2. RNA-Seq results

Millions of sequencing reads were obtained through RNA-Seq. For each sample, around 24 million reads, aligned to 67k Unigenes, which corresponded to an average of approximately 350 reads per Unigene. However, a large percentage of the reads aligned to a few genes that were highly expressed (Figure 6). Consequently, most of the genes were represented by only a few reads. RNA-Seq power lies in the quantification of all mRNA in a particular sample, but this also makes detection and quantification of genes expressed at low level more difficult. As shown in this study, a few genes were highly expressed, which made the detection of most genes more difficult. Consequently, with the expression of most genes relying on only a few reads, the DE analysis could have been greatly affected. The choice of sequencing depth is therefore of utmost importance and no less than the one used in this study is recommended for the sequencing of wheat leaf transcriptome.

Analysis of the similarity between samples (Figure 7) showed that the samples were grouped into three main clusters, each representing a time point. The transcriptome profile of one genotype at a particular time point was thus more similar to that of a different genotype at the same time point, than that of the same genotype at a different time point. This result indicates that the main factor influencing gene expression was time. This was also shown by the high number of DE genes between time points, independent of genotype (see Results section II.2). Many genes may be DE through time because of responses to environmental conditions or normal plant development, rather than because of responses to rust. However, the grouping of genotypes into three sub-clusters at 2 and 5 DPI (Figure 7), but not at 0 DPI, indicates that a factor at these time points allowed the differentiation between the different genotypes, when at 0 DPI they seemed undistinguishable. The proximity of Col-NS765 and Col-NS766, compared to Col, strongly suggests that this factor was the rust, and that the resistance response was triggered in the resistant lines at these time points.

Finally, the higher number of DE genes between Col and Col-NS765, than between Col and Col-NS766, indicates that the transcriptome profile of Col is more similar to that of Col-NS766 than that of Col-NS765. A possible explanation could be that more of the CTH-NS genetic background remains in the backcrossed Col-NS765 line than in the Col-NS766 line.

II.3. Genes of interest

II.3.1. Loci of resistance on chromosomes 7AL and 6A

It is likely that the genes of interest that were DE between Col and Col-NS at 0 DPI were DE because they were derived from the CTH-NS genetic background, and were selected along with the resistance phenotype through five backcrossed generations in both of the two independent Col-NS lines. This class of genes was enriched for those mapping to chromosomes 6AL, 6AS and 7AL (Figure 10B), suggesting that these genes are part of genomic regions that are polymorphic between Col and Col-NS. Moreover, many of these genes were only expressed in Col or Col-NS, indicating that they may be absent in one line and present in the other. This was not surprising for 7AL as the major locus conferring resistance to stem rust in Col-NS was mapped to this chromosome arm. DE genes located on this chromosome are, therefore, candidates for the initiation of the resistance response.

Interestingly, even more genes were found DE on chromosome 6A. This is consistent with the mapping data that showed a high degree of polymorphism between Col and Col-NS766 on chromosome 6A (see chapter three). However, in this experiment, both Col-NS765 and Col-NS766 were used for detecting genes that were DE between resistant and susceptible lines. This indicates that Col-NS765 also possesses a large genomic region on 6A that was introgressed from CTH-NS. Because Col-NS765 and Col-NS766 are two independent backcrossed lines (BC_5F_4), the chance that this region was randomly selected in both lines is low. Therefore, this region may have been actively selected because it contains genes that contribute to the resistance against stem rust. These genes would be dispensable, since genetic mapping showed that the 7AL locus was the main locus conferring resistance. Nevertheless, they may still affect the resistance and could be the cause of the difficulty encountered to precisely map the 7AL locus. This would be in agreement with the multigenic resistance hypothesis (see chapter three).

II.3.2. Origin genes (0 DPI and on 7L)

Among the genes of interest, most that were DE at 0 DPI were DE between resistant and susceptible lines at all time points (Figure 9). In addition, most of them were constitutively expressed (if expressed) in all genotypes. Although not regulated by rust inoculation, some of these genes could be involved in the pathogen recognition and the initiation of the defence response, especially genes found on 7AL, since the locus conferring resistance is on this chromosome. Genes that were DE between Col and Col-NS at 0 DPI and located on 7AL, are discussed below. Genes on 7BL and 7DL were also considered, in case they corresponded to 7AL homoeologues that were missing in the Unigene or GSS databases. Genes that were more expressed in Col-NS than in Col were considered as putative resistance genes, whereas those that were more expressed in Col were considered as putative susceptibility genes, that is, genes that could promote susceptibility.

Concerning the putative susceptibility genes, 10 genes were found, of which five were annotated (Table 3, Origin Down). Based on the annotation, Ta.42512, which codes for a member of the very large family of cytochrome P450 (CYP), suggests an implication in plant defence. These enzymes are known to be involved in many biosynthetic reactions, including in plant defence pathways (Nelson and Werck-Reichhart, 2011). Ta.42512 is most closely related to the CYP71 clan, the largest CYP clan, which includes CYP71B15 found to be involved in the biosynthesis of the phytoalexin Camalexin in *Arabidopsis* (Schuhegger et al., 2006). Interestingly, this gene was highly expressed in Col at 2 and 5 DPI. It is unknown if this gene can promote susceptibility.

Cysteine-rich receptor-like kinases: R-genes candidates

Concerning the putative resistance genes, 19 were found, of which six were annotated and possibly involved in the resistance (Table 3, Origin Up). These genes (Ta.75386, Ta.90783, Ta.90847, Ta.95070, Ta.104804 and Ta.109657) all coded for cysteine-rich receptor-like kinases (CRKs). CRKs comprise a large subfamily of receptor-like kinases (RLKs), and have been shown in *Arabidopsis*, to be induced by pathogen attack and salicylic acid (SA) (Du and Chen, 2000; Chen et al., 2004; Acharya et al., 2007; Wrzaczek et al., 2010). Some CRKs have been shown to activate cell death and to increase resistance against pathogens (Chen et al., 2003, 2004; Acharya et al., 2007). Conversely, in some studies, loss-of-function mutations and silencing of CRKs enhanced resistance to pathogens (Ederli et al., 2011; Rayapuram et al., 2012). Two of

the selected CRKs (Ta.104804 and Ta.95070) possess a jacalin-like lectin domain. Proteins with such domain have been shown to be associated with stress response and fungal resistance, such as TaJRLL1 in wheat (Xiang et al., 2011). These genes seem therefore good candidates for being involved in the induction of the defence response.

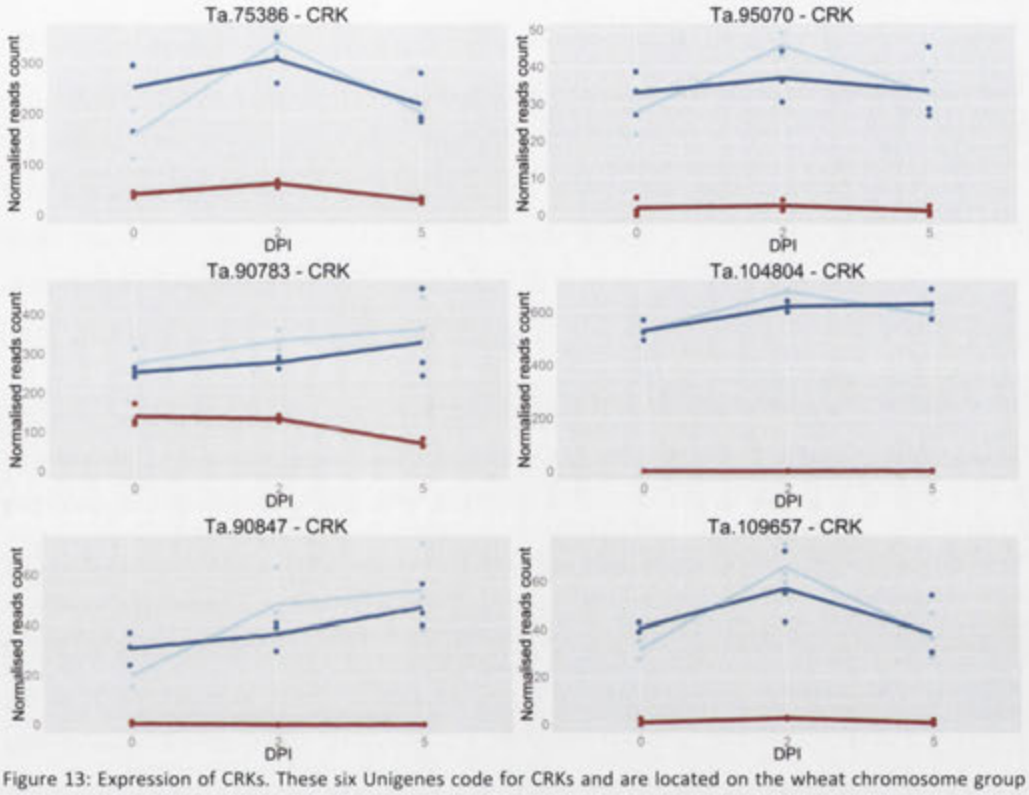


Figure 13: Expression of CRKs. These six Unigenes code for CRKs and are located on the wheat chromosome group 7L. They are more expressed in Col-NS than in Col at any time point, and some are not expressed in Col. For each Unigene, the DESeq-normalised reads count is reported for each sample (dots) at the different time points. Lines represent the biological replicates average. Colours represent genotypes: Col (red), Col-NS765 (pale blue), Col-NS766 (dark blue).

Moreover, two of these genes were mapped and found to be linked to stem rust resistance (see chapter three figure 12); Ta.104804, which was previously found to be located on a GSS contig linked to the resistance, and Ta.90783. Several single-nucleotide polymorphisms (SNPs) between Col and Col-NS were found in Ta.90783, by comparison of the RNA-Seq reads from the Col and Col-NS genotypes. Although Ta.104804 (marker 7AL_4519254_7775) showed some recombination events with the phenotype or the cluster of co-segregating markers (21 and 11 out of 150 $F_{2:3}$ families, respectively), Ta.90783 (marker 7AL_4487195) co-segregated with the cluster. As this cluster corresponds to the most probable location for the 7AL locus, Ta.90783 is a good candidate for being involved in the resistance.

Ta.90783, which is located on 7AL, was found highly similar to Ta.90847, which is located on 7DL (BLASTN: 94 % identities, e-value= 0). Concerning Ta.90847, only reads from Col-NS aligned to a small region of less than 200 bp. In contrast, reads from any genotype aligned to the full length of Ta.90783, except at the corresponding regions of Ta.90847 with aligned Col-NS reads, for which only reads from Col aligned. For this region, SNPs were found between Col and Col-NS, which probably led to the misalignment of reads from Col-NS to Ta.90847, when they should have aligned to Ta.90783. Consequently, Ta.90847 is most likely the 7DL homoeologue of Ta.90783 on 7AL and is probably not expressed as the alignment of reads to this gene could be an artefact. As Ta.90783 is one of the best candidates for being the gene conferring resistance in Col-NS, Ta.90847, which is not expressed in the Col background, would be a possible candidate for being the CTH 7DL suppressor, in a similar manner to the rye R-gene, *Pm8*, that is suppressed by the orthologous wheat gene *Pm3* (McIntosh et al., 2011; Hurni et al., 2013).

Apart from Ta.90847, the other CRKs in the Origin/Up category are not closely similar in DNA sequence, but their putative orthologues in rice all belong to the same small sub-family of receptor-like cytoplasmic kinases, (RLCK)-OS2 (Dardick et al., 2006; Jung et al., 2010). This may indicate some functional relatedness between these wheat CRKs. This sub-family seems, however, much bigger in wheat as the best rice hit for 270 Unigenes were one of the RLCK-OS2 members. Interestingly, the rice RLCK-OS2 family contains many members with strong sequence similarities to *Rpg1*, which encodes a RLK with two tandem protein kinase domains and confers broad-spectrum resistance to *Pgt* in barley (Brueggeman et al., 2002; Dardick and Ronald, 2006). In addition, Ta.90783 and Ta.104804 have good similarities with RPG1. This gene was found constitutively expressed in barley, and the RPG1 protein was found to be phosphorylated within five min after inoculation with virulent urediniospores (Rostoks et al., 2004; Nirmala et al., 2010). Two effectors present in the stem rust urediniospores were found to induce RPG1 phosphorylation as well as HR (Nirmala et al., 2011). The similarities between the Rpg1-mediated defence response and the resistance response in Col-NS, which are both activated at an early stage and before the formation of haustoria, reinforce the possibility that the above CRKs, which are part of the same sub-family as RPG1, may be involved in the activation of the resistance response.

Finally, RLKs that are involved in disease resistance seem to have undergone many duplication events, in rice and *Arabidopsis*, and are often found in tandem clusters (Shiu, 2004).

Here, it is unclear if these CRKs are clustered, but they all map to 7AL, and their enrichment in the category of putative resistance genes, compared to other categories, supports the involvement of these genes in the resistance. On the other hand, the 7AL resistance gene may be either not represented in the Unigene and GSS databases, or not be DE between Col and Col-NS, and thus would be undetectable in this study.

II.3.3. Constitutive genes (not DE through time)

153 (43 %) genes of interest were categorized as constitutive, that is, genes that were not DE through time in any genotype (Figure 14). Although all could be involved in the resistance, most are expected to be on polymorphic regions between Col and Col-NS (i.e. regions of Col replaced by the corresponding regions in CTH). 92 were more expressed in Col-NS than in Col, of which 13 were annotated. Three genes were of particular interest, based on their annotation (Table 3, Constitutive Up); Ta.155345 and Ta.111478, which both encode CRKs and, like the ones found in the Origin category, are also part of the RLCK-OS2 sub-family, and Ta.83884 which encodes a TRANSPORTIN 1-LIKE (TRN1-LIKE) protein. In *Arabidopsis*, AtTRN1 is a nuclear import receptor for AtGRP7 (Ziemienowicz et al., 2003), which is a glycine-rich RNA-binding protein demonstrated to confer resistance against the bacteria *P. Syringae* DC3000 and *P. carotovorum*, as well as against *tobacco mosaic virus* (Fu et al., 2007; Lee et al., 2012).

61 genes were less expressed in Col-NS than in Col, of which five code for disease resistance proteins (Table 3, Constitutive Down), including three nucleotide-binding leucine-rich repeat (NB-LRR) proteins resembling RGA1 (Ta.93375, Ta.18874 and Ta.174313), one resembling RPM1 (Ta.12047) and one with lower homology to RGA1 (Ta.39158). These genes were not expressed or expressed at low level in Col-NS (Figure 14). This finding is surprising as NB-LRRs are usually associated with pathogen recognition and defence, rather than with susceptibility as this result would imply (Marone et al., 2013). NB-LRR type proteins can mediate susceptibility, for example *lov1*, in *Arabidopsis*, which encodes a NB-LRR and confers sensitivity to the necrotrophic fungus *C. victoriae* due to the production by the pathogen of the toxin victorin (Lorang et al., 2007; Sweat et al., 2008). However, the promotion of susceptibility by NB-LRR was never shown with biotrophic pathogens.

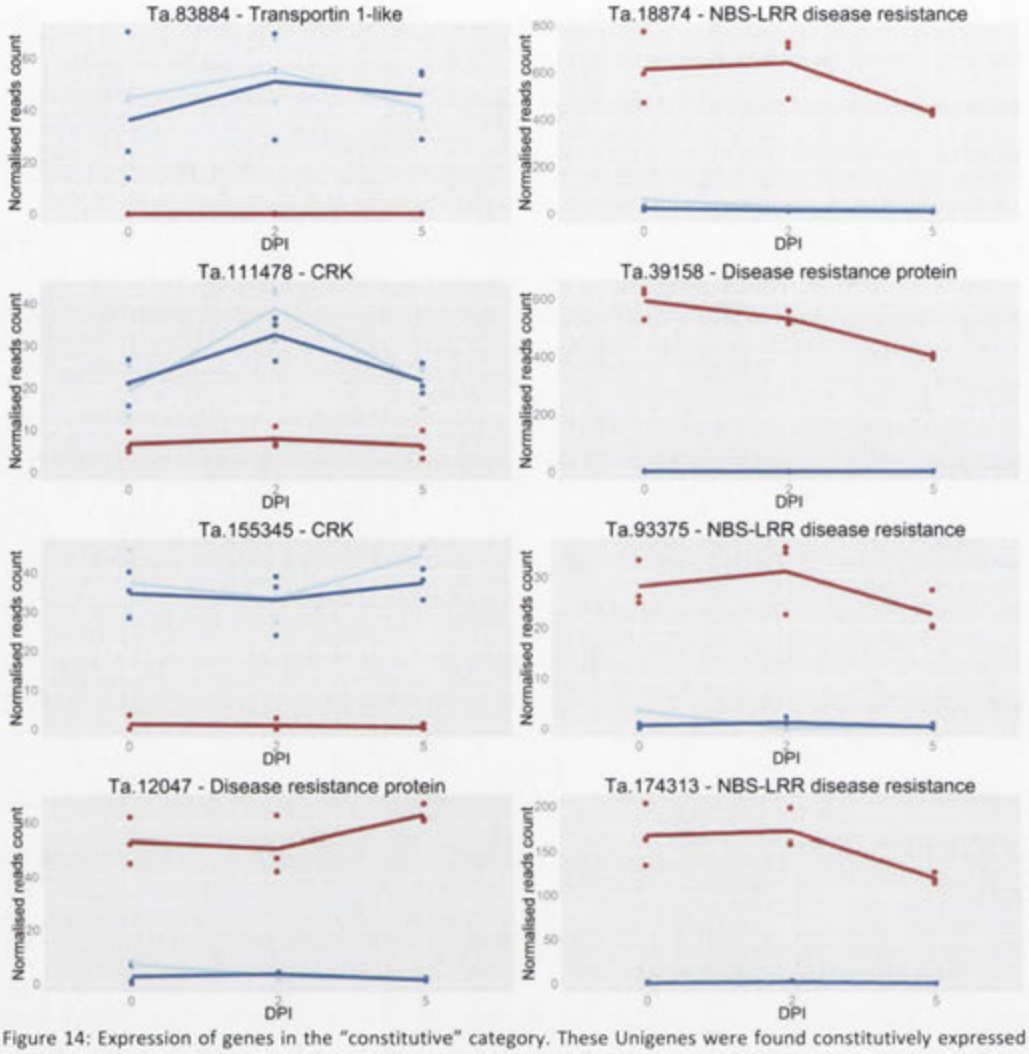


Figure 14: Expression of genes in the "constitutive" category. These Unigenes were found constitutively expressed in all genotypes. For each Unigene, the DESeq-normalised reads count is reported for each sample (dots) at the different time points. Lines represent the biological replicates average. Colours represent genotypes: Col (red), Col-NS765 (pale blue), Col-NS766 (dark blue).

II.3.4. Response Genes (2 and 5 DPI)

An increase of 29 and 41 % genes DE between Col and Col-NS were found at 2 and 5 DPI, respectively, compared to 0 DPI (Figure 9). This transcriptomic change could be due to the rust infection and the implementation of the defence response in the resistant lines. Consequently, these genes could be involved in the resistance response. Depending on the time point at which they were DE, genes were classified as early response, if only DE at 2 DPI, late response, if only DE at 5 DPI, and both responses, if DE at both 2 and 5 DPI.

Early response (2 DPI)

38 genes were more expressed in Col-NS than in Col, of which seven were annotated (Table 3, Early Up). All these genes could be involved in the defence response, notably Ta.1164 (Serine/threonine-protein kinase) and Ta.45840 (Calmodulin-binding protein), which could be implicated in signal transduction. However, their expression patterns were very similar in Col and Col-NS, which could suggest that these genes may not be the ones responsible for differentiation between resistance and susceptibility (Figure 15). Likewise, many of the genes of interest also showed the same pattern of expression in all genotypes. Nevertheless, the differential expression of these genes could have a quantitative effect on the resistance response. Ta.25495 and Ta.7579, which encode a deoxyuridine 5-triphosphate nucleotidohydrolase and a transporter, respectively, may have different patterns of expression in Col and Col-NS, especially the latter, which was down-regulated in Col but not in Col-NS, at 2 DPI (Figure 15).

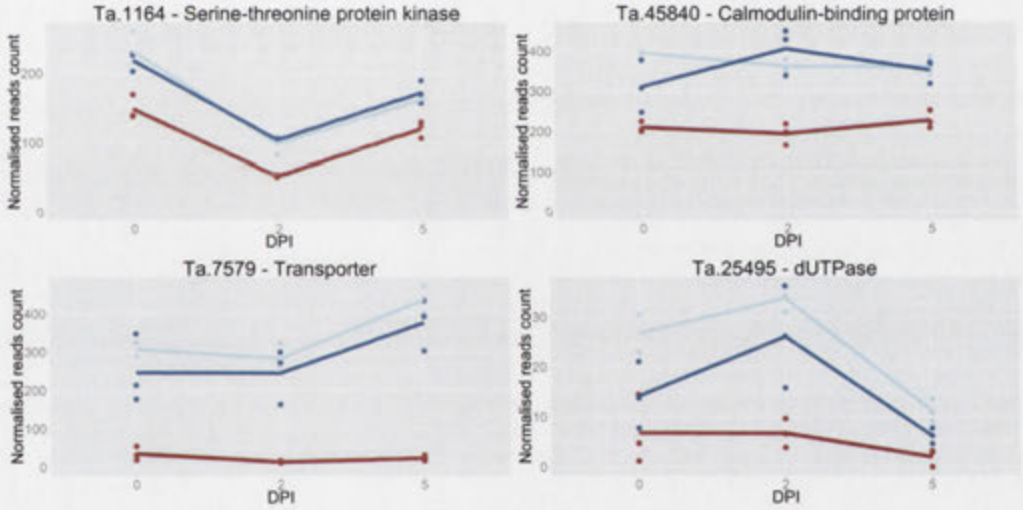


Figure 15: Expression of selected genes in the "early up" category. These Unigenes were found more expressed in Col-NS than in Col at 2 DPI. For each Unigene, the DESeq-normalised reads count is reported for each sample (dots) at the different time points. Lines represent the biological replicates average. Colours represent genotypes: Col (red), Col-NS765 (pale blue), Col-NS766 (dark blue).

25 genes were less expressed in Col-NS than in Col. 12 were annotated (Table 3, Early Down), of which four were of particular interest. Ta.13785 codes for a xylanase inhibitor. Xylanases are used by pathogens to degrade the plant cell wall and inhibitors are produced by the plant to counteract their effects (Dornez et al., 2010). The expression of a xylanase inhibitor in rice has been reported to increase in response to jasmonic acid (JA) (Xin et al., 2014). However, this gene was down-regulated in Col-NS at 2 DPI, and in all genotypes at 5 DPI

(Figure 16). It is unclear why this gene was highly expressed at 0 DPI in all genotypes. Perhaps, it has been activated during the time the plants were left to dry, after rust inoculation and before the sample collection for 0 DPI (approximately 30 min). Two other genes, Ta.39365 and Ta.105098, presented a similar expression pattern as Ta.13785 (Figure 16), and both code for linalool synthase. Linalool is a monoterpene found to have a role in defence and the expression of linalool synthase was found to be up-regulated by JA (Schie et al., 2007; Taniguchi et al., 2013). The sudden decrease in the expression of these three genes, in the resistant genotypes, could reflect an antagonist pathway of JA such as salicylic acid (SA). Finally, Ta.74923 code for an endo-1,3;1,4-beta-D-glucanase-like, which could be involved in cell wall turnover in the Poaceae family (Høj and Fincher, 1995). The down-regulation in Col-NS could indicate a slowdown of cell elongation during the defence phase (Figure 16).

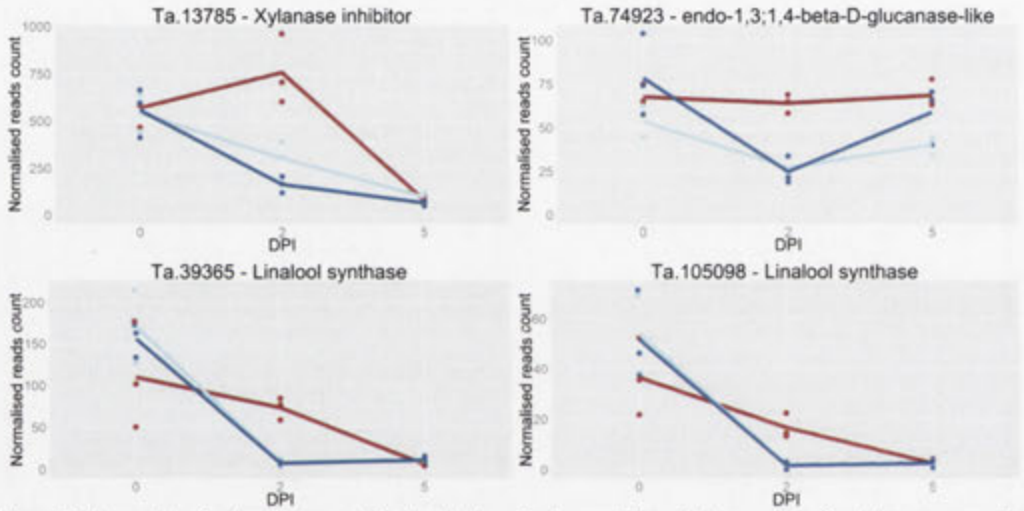


Figure 16: Expression of selected genes in the "early down" category. These Unigenes were found less expressed in Col-NS than in Col at 2 DPI. For each Unigene, the DESeq-normalised reads count is reported for each sample (dots) at the different time points. Lines represent the biological replicates average. Colours represent genotypes: Col (red), Col-NS765 (pale blue), Col-NS766 (dark blue).

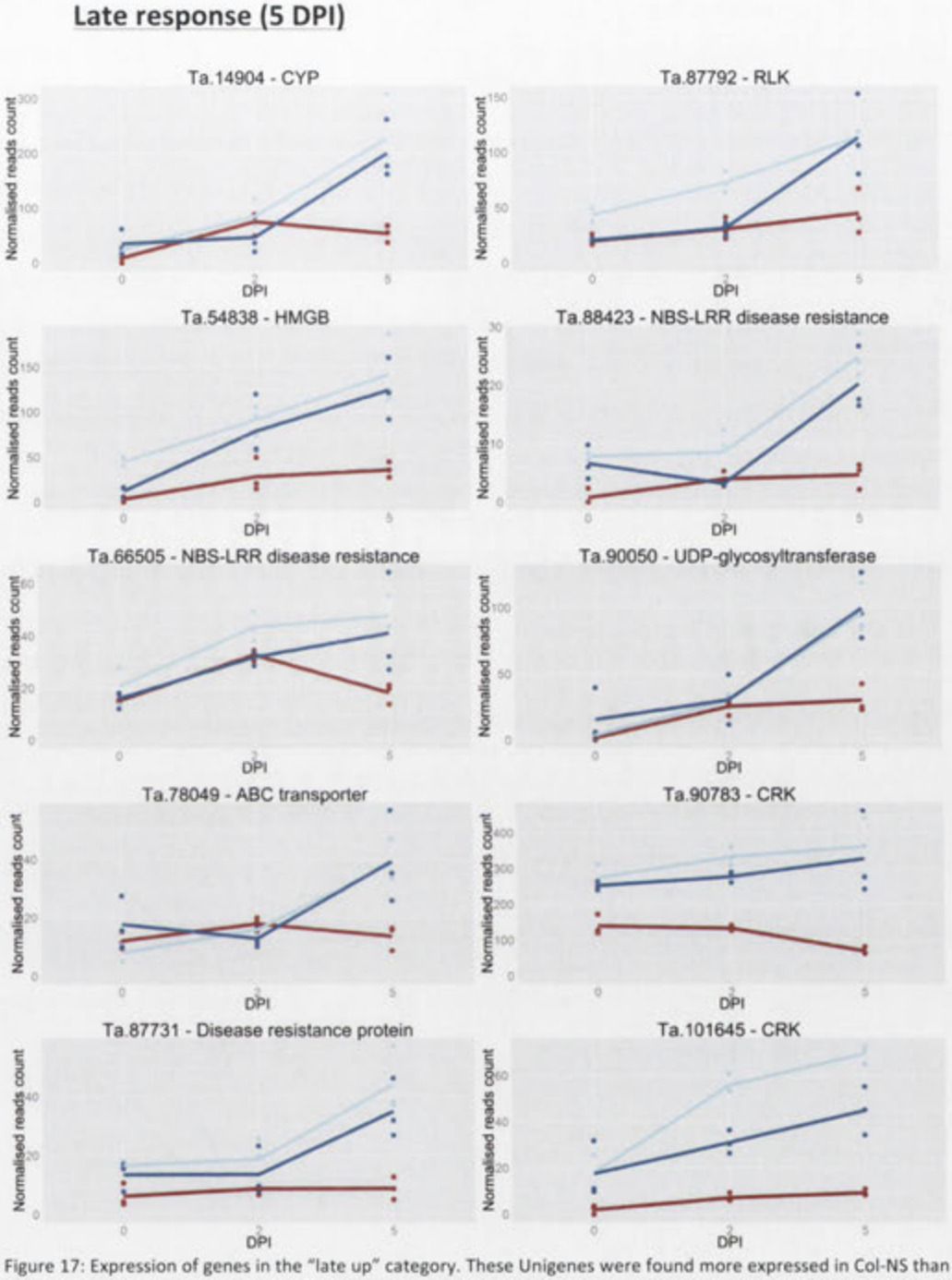


Figure 17: Expression of genes in the "late up" category. These Unigenes were found more expressed in Col-NS than in Col at 5 DPI. For each Unigene, the DESeq-normalised reads count is reported for each sample (dots) at the different time points. Lines represent the biological replicates average. Colours represent genotypes: Col (red), Col-NS765 (pale blue), Col-NS766 (dark blue).

46 genes were more expressed in Col-NS than in Col. 26 were annotated (Table 3, Late Up), of which 15 were of particular interest (Figure 17). Ta.66505, Ta.87731 and Ta.88423 code

for disease-resistance proteins. Ta.75386, Ta.90783, Ta.101645 and Ta.103989 code for CRKs. The first two were already found as putative resistance genes. Ta.87792 and Ta.72300 code for RLKs. Ta.89279 code for a serine/threonine-protein kinase receptor. Ta.14904 code for a CYP. Ta.78049 code for an ABC transporter, which could be involved in many processes from detoxification to pathogen resistance (Kang, 2011). Similarly, Ta.77626 is part of the multidrug and toxic compound extrusion (MATE) family, responsible for the transport of various compounds (Conte and Lloyd, 2011). Ta.90050 code for UDP-glycosyltransferase, which could be involved in the SA-dependent defence response (Langlois-Meurinne, 2005; Park et al., 2011). Finally, Ta.54838 code for a member of the high mobility group B (HMBG) proteins family, which is known to be involved in various cellular processes such as transcription and replication, and is regulated by various abiotic stresses (Kwak et al., 2006).

34 genes were less expressed in Col-NS than in Col. 16 were annotated (Table 3, Late Up), of which four were of interest (Figure 18). Ta.55220 and Ta.14588 code for polygalacturonase inhibitors. Polygalacturonases are known to be secreted by pathogens to degrade the plant cell wall and their inhibition is part of the plant defence (Lagaert et al., 2009). They are, however, down-regulated in the resistant lines. Ta.11076 codes for an UDP-glycosyltransferase and Ta.195306 for a CYP. They were both up-regulated in Col but not in Col-NS.

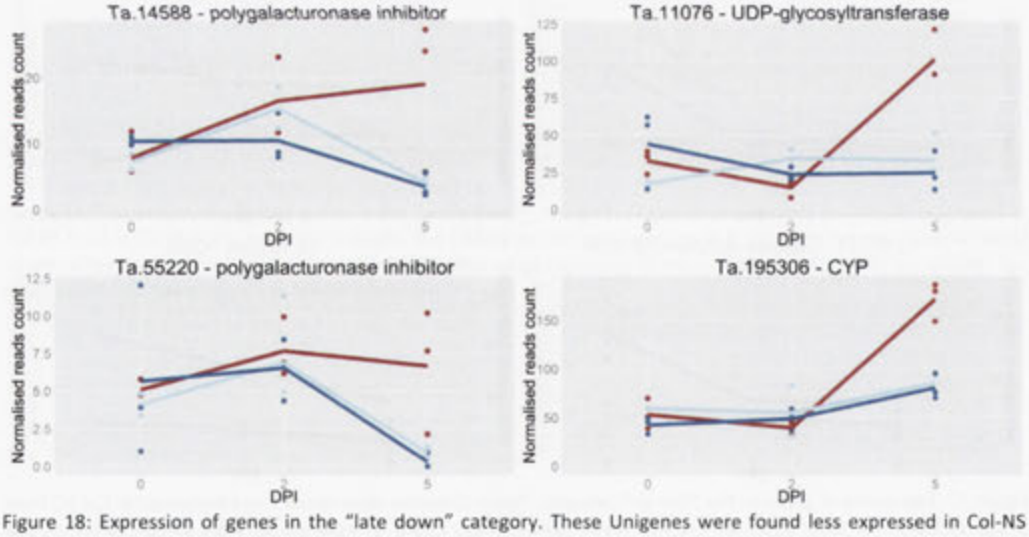


Figure 18: Expression of genes in the "late down" category. These Unigenes were found less expressed in Col-NS than in Col at 5 DPI. For each Unigene, the DESeq-normalised reads count is reported for each sample (dots) at the different time points. Lines represent the biological replicates average. Colours represent genotypes: Col (red), Col-NS765 (pale blue), Col-NS766 (dark blue).

Four extra genes were found part of the late response category and were less expressed in Col-NS than in Col. However, these genes were not from the plant but from the

pathogen (Table 3, Pathogen). Ta.68873 and Ta.68881 both code for 40S ribosomal proteins, Ta.68870 for a 60S ribosomal protein and Ta.114604 for an elongation factor Tu. They were not expressed in any genotype at 0 DPI and not or almost not at 2 DPI. However, at 5 DPI, they were highly expressed in Col, whereas they remained expressed at low level in Col-NS. Using BLASTX against the NCBI nr database, these genes showed high similarities with genes in stem rust and various fungi. They were integrated to the wheat Unigene database from a cDNA library of a compatible interaction between stripe rust (*P. striiformis*) and wheat. Unsurprisingly, these results confirm the restriction of pathogen growth in the Col-NS genotype.

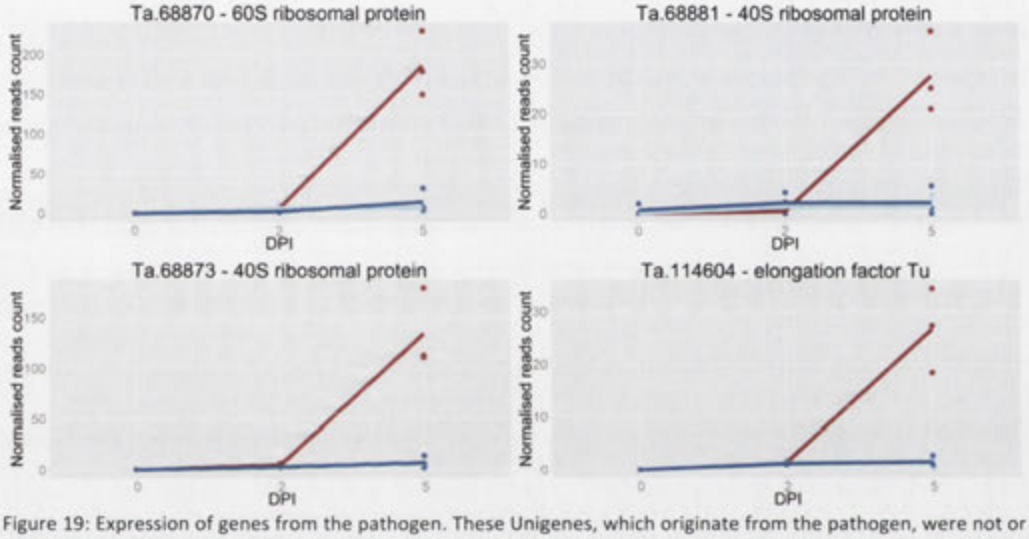


Figure 19: Expression of genes from the pathogen. These Unigenes, which originate from the pathogen, were not or lowly expressed at 0 and 2 DPI, and were highly up-regulated in Col at 5 DPI. For each Unigene, the DESeq-normalised reads count is reported for each sample (dots) at the different time points. Lines represent the biological replicates average. Colours represent genotypes: Col (red), Col-NS765 (pale blue), Col-NS766 (dark blue).

Both Early and Late responses (2 and 5 DPI)

Seven and nine genes, which could be involved in both early and late resistance responses, were more and less expressed in Col-NS, compared to Col, respectively. Five were annotated (Table 3, Both), of which two could be of interest for the promotion of susceptibility: Ta.87395, which code for a LRR receptor-like serine/threonine-protein kinase, and Ta.42512, which code for the CYP already discussed in the putative susceptible genes. The DE for these genes could be due to differing alleles in Col and Col-NS.

II.4. Hormonal pathways

The study of genes that were DE between Col and Col-NS led to the discovery of many genes potentially involved in the resistance. However, the implication of hormonal pathways that are usually involved in plant defence, such as the SA and JA pathways, is unclear. SA and JA are often described as antagonists because SA has a role in resistance to biotrophic pathogens whereas JA has a role in defence against herbivores and necrotrophs (An and Mou, 2011; Wasternack and Hause, 2013). Although these pathways have been thoroughly studied in *Arabidopsis*, most of the genes involved in these pathways have not been investigated in other plants and evidences indicate that the *Arabidopsis* pathways could be different from that of other plants (Matsushita et al., 2013). Nevertheless, pathogenesis-related (PR) genes are often seen as the outcome of these pathways. PR1, PR2 and PR5, are usually associated with the SA pathway whereas PR3 and PR4 are usually associated with the JA pathway. To investigate these PR genes directly, the putative corresponding Unigenes were retrieved (Table 4) and their expression was analysed in the RNA-seq data.

All five PR genes were found induced in all genotypes at 2 and 5 DPI (Figure 20). In no case there was a significant difference between genotypes. These results strongly suggest that both SA and JA pathways have been activated in all three genotypes. However, the lack of differential expression between genotypes indicates that the defence response in Col-NS is independent of these pathways, or that it is downstream or in parallel of their activation. Interestingly, most of these genes were found highly expressed in one of the samples for Col-NS766 at 0 DPI, but not or lowly expressed in all the other samples at the same time point. Many genes were in fact more expressed in this sample than in the two other biological replicates, notably genes with annotations related to plant defence (data not shown). This suggests that the defence responses may have already been activated in this particular sample, which was thus considered an outlier. However, this did probably not affect much the transcriptome analysis.

Family	Name	Accession	Unigene	Reference
PR1	Pathogenesis-related protein 1-3	HQ541963	Ta.102727	(Molina et al., 1999)
PR2	Beta-1,3-glucanase	DQ090946	Ta.35829	(Liu et al., 2010)
PR5	Thaumatin-like	X58394	Ta.56897	(Singh et al., 2012)
PR3	Chitinase 1	AB029934	Ta.70604	Unpublished
PR4	Pathogenesis-related protein 4	AF092123	Ta.58603	Unpublished

Table 4: Details of putative PR genes.

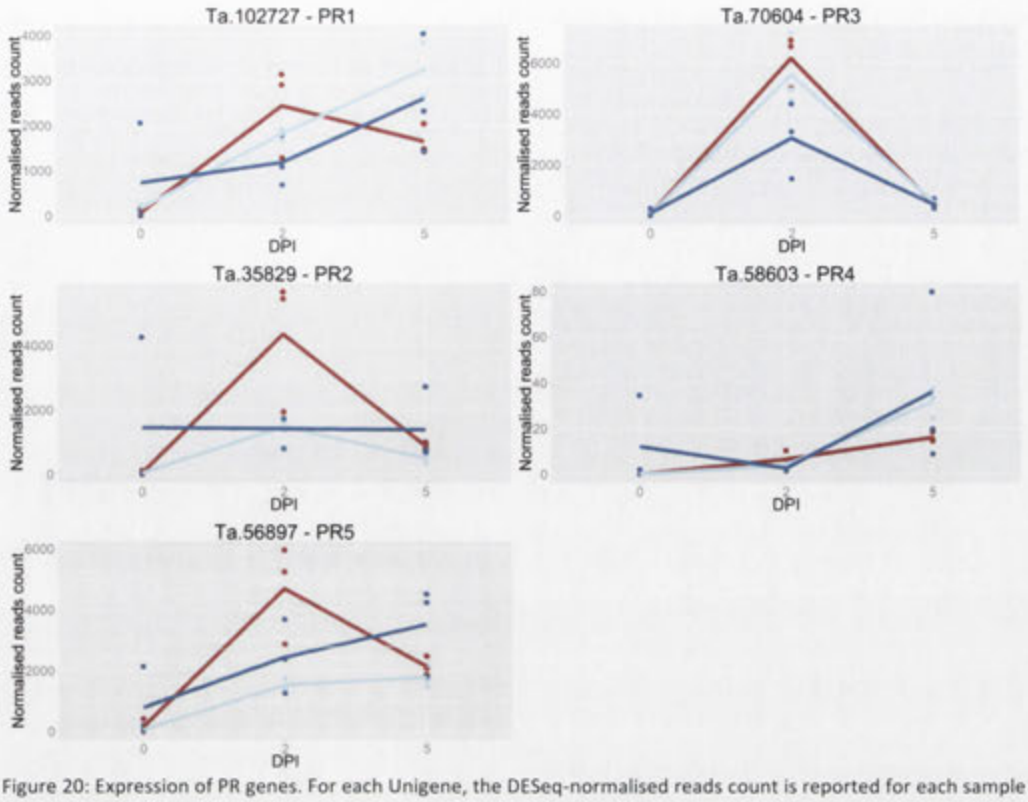


Figure 20: Expression of PR genes. For each Unigene, the DESeq-normalised reads count is reported for each sample (dots) at the different time points. Lines represent the biological replicates average. Colours represent genotypes: Col (red), Col-NS765 (pale blue), Col-NS766 (dark blue).

II.5. Mock-inoculation control

This study allowed detecting many genes that were DE between the susceptible and resistant lines, indicating that of all detected genes, these are the most likely for being involved in the defence mechanism. However, it is unclear if these transcriptomic changes, notably for the genes found DE at 2 and 5 DPI, were due to the rust infection or the inoculation conditions. In order to address this concern, preliminary RT-qPCR experiments were performed for some genes, using the samples used in RNA-Seq, as well as mock-inoculated samples. As expected, results from RT-qPCR followed the same trend as that of RNA-Seq (Figure Sup. 2). For some genes, their expression was different in rust- and mock-inoculated samples. This was the case for Ta.48480, which code for a PAL protein. This gene was highly down-regulated in rust-inoculated samples at 2 and 5 DPI, similarly to the RNA-Seq results, whereas its expression level remained relatively constant in mock-inoculated samples (Figure 21A). This suggests that its expression was related to the rust infection. Other genes were regulated the same way in rust- and mock-inoculated samples. For instance, Ta.88423, which code for a disease re-

sistance protein and was found putatively part of the late defence response, was also up-regulated in the mock-inoculated Col-NS766 at 5 DPI and significantly more expressed than in Col (Figure 21B). Therefore, the inoculation conditions, along with the genetic background, and not the rust infection, seemed to be the cause for this transcriptomic change.

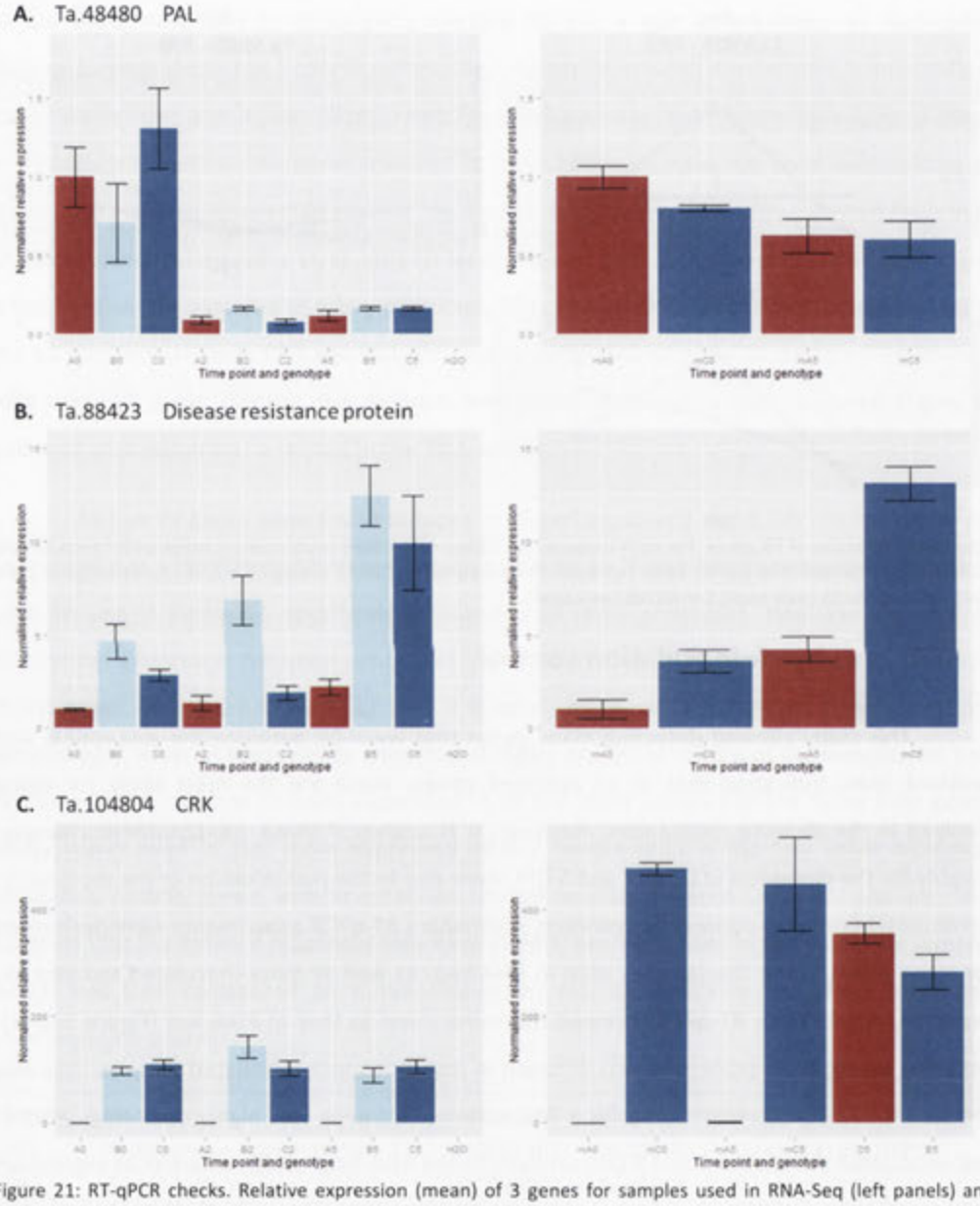


Figure 21: RT-qPCR checks. Relative expression (mean) of 3 genes for samples used in RNA-Seq (left panels) and mock-inoculated samples (right panels). Letters represent genotypes (A: Col, B: Col-NS765, C: Col-NS766, D: CTH-K, E: CTH-NS1) and numbers time points (0: 0 DPI, 2: 2 DPI, 5: 5 DPI). Mock-inoculated samples are indicated by "m". For each sample, the gene expression was normalised to the corresponding Col at 0 DPI (A0 or mA0), thus level of expression between rust- and mock-inoculated cannot be directly compared. Errors bars represent the standard error of the mean for triplicates (left panels) or duplicates (right panels). The relative expression level was calculated by comparing the target gene with the reference Unigene Ta.2291 (ADP-ribosylation factor).

As demonstrated by the RT-qPCR results, the variation in the expression of some of the genes of interest identified in this study could be unrelated to the rust infection. RT-qPCR can be used to investigate this point for individual genes of interest. The aim of this study was to identify genes that were DE between the susceptible (Col) and two independent resistant (Col-NS) lines. The use of two independent resistant lines displaying the same genetic resistance is a powerful method for focusing on gene expression differences relevant to this specific mechanism of resistance. The experimental design used here also allowed the identification of genes that were not DE through time, but probably differ in expression between genotypes because they represent alleles from the different genetic backgrounds. For instance, Ta.104804, which codes for a CRK and was categorized as a putative resistance gene, was found, through RT-qPCR, to be similarly expressed in the susceptible line, CTH-K, and the resistant lines, Col-NS766 and CTH-NS1, whereas it was never expressed in Col (Figure 21C), indicating that this gene was inherited from CTH in Col-NS. This was demonstrated directly for Ta.90783 (CRK) for which a SNP marker (7AL_4487195) was showing the same allele in Col-NS and CTH, but was different in Col (see chapter three).

III. Summary

The observation, at a microscopic level, of the rust infection in the susceptible line, Col, and the resistant lines, Col-NS765 and Col-NS766, gave a better understanding of the defence mechanism. This defence consists in the death of cells that are adjacent to an infection site, that is, where the appressorium was formed. This indicates that the resistance in Col-NS is part of ETI, since cell death is usually the sign of HR, the ETI landmark. However, an enhanced response of PTI cannot be excluded, especially since the cell death was observed at the epidermal level and not at the mesophyll level, as it is usually the case in ETI (Catanzariti et al., 2007). In addition, more *Pgt* races need to be tested on Col and Col-NS, in order to determine if the resistance in Col-NS is race-specific, which could also be indicative of ETI. In some occasions, the fungus could still spread into the resistant lines, but its growth was slower than in Col, possibly due to a secondary resistance mechanism. Interestingly, some level of resistance was also observed in Col, and was similar than in Col-NS. The same ETI response may be involved in both lines but may be more important in Col-NS.

Through RNA-seq, a total of 353 genes were found DE between Col and both Col-NS765 and Col-NS766. Overall, there was a good difference between resistant and susceptible

lines. However, in many cases, the pattern of gene expression through time was very similar across genotypes, but the level of expression varied between genotypes. Although the difference may be statistically significant, it may not be biologically relevant, as the regulation appears to be similar. In addition, genes were sometimes found DE between genotypes at a certain time point but not at another, even though these genes were not DE through time in any genotype. This was mainly due to the slight up- and down-regulation of a particular gene across the genotypes. Again, these changes may not be meaningful at a biological level. Unsurprisingly, around two third of the Unigene were lacking annotations, and except for their expression pattern, no clues were given about their role in the plant. However, the aim of this study was to obtain a first insight of the genes involved in the resistance against stem rust. Further studies need to be conducted in order to determine the role and implication of these genes.

In agreement with the mapping data described in chapters two and three, the chromosomal location of the genes of interest indicated the chromosome 7AL as a source of genes that initiate the defence response. Interestingly, genes on chromosome 6A could also be involved in this process, as this chromosome was selected in two independent backcrosses. This result supports the hypothesis of multiple segregating resistance genes in the mapping population (see chapter three).

Many of the genes of interest were always DE and often in a constitutive manner, indicating that their expression was not affected by the rust inoculation. However, these genes could be regulated at the post-translational level, and could still have a role in defence. Moreover, some of these genes, notably those coding for CRKs, are good candidates for being resistance genes involved in initiating the defence response. Therefore, it would be of great interest to undertake functional analyses of these genes. Concerning the defence response, histological observations showed that the first signs of the resistance affecting the rust development occurred at 2 DPI. However, fewer genes potentially involved in the early defence response at 2 DPI were found, than genes potentially involved in the late defence response at 5 DPI. As discussed above, since the resistance response at 2 DPI was limited to the cells adjacent to an infection site, the expression of the genes involved in the early resistance response may have been diluted, thus their differential expression was undetectable. At 5 DPI, the resistance may be more global, due to the expansion of the fungus in the plant, or due to the onset of a secondary defence mechanism, such as SAR.

MATERIALS AND METHODS

Stem rust inoculation

For histological and transcriptome analyses, plants were inoculated with wheat stem rust *Puccinia graminis* f. sp. *tritici*, race #313 (pathotype 34-1,2,3,5,6,7), at the Plant Breeding Institute, University of Sydney. Seedlings were grown in growth chamber (20-22°C, natural light cycle - max 170 µmol/m^2/s) until 2-leaf stage (≈2 weeks). Inoculation was done by spraying the plants with spores suspended in oil. Plants were let to dry for 30 min and incubated for 2 days in incubation cabinets (20-22°C, high humidity, natural light cycle - max 85 µmol/m^2/s). Plants were finally moved back to growth chamber at lower temperature (17-19°C) and scored for stem rust resistance at 12 DPI.

Histological observations

Histological observations were done on the pool of first-leaf from two plants, using Zeiss Axio Imager.M1 microscope with the AxioVs40 software (v4.8.20). WGA-FITC staining was done following a modified protocol described by Ayliffe et al. (2011). Briefly, leaf tissue were immersed in 1 M KOH and incubated at 37°C for 2 days. KOH solution was replaced during incubation. Tissue was then neutralized in 50 mM Tris, pH 7.0, and stained with WGA-FITC at 20 µg/ml (Sigma-Aldrich, St. Louis). Fungal structures and autofluorescence were observed under blue (excitation 470-440 nm, emission 525-550 nm) and UV (excitation 365 nm, emission ≥ 420) light, respectively. Trypan blue staining was done following the method describe by Desmond et al. (2008). These experiments were performed twice, with similar observations on both occasions.

Transcriptome analyses

For each sample, the first-leaf from two plants were pooled and quickly frozen in liquid nitrogen. Total RNA was extracted using the RNeasy Plant Mini Extraction Kit and RNA integrity was checked on 1.4 % agarose gel and on the Agilent® 2100 Bioanalyzer. Following a poly(A) isolation step performed by the sequencing company, mRNA samples were sequenced using Illumina® HiSeq 2000 for 100bp single-end and an expected yield of 30 million reads. Reads were aligned against the wheat (*Triticum aestivum*) Unigene database build #62 (http://www.ncbi.nlm.nih.gov/unigene), comprising 158,028 transcripts, using the BioKanga software (http://code.google.com/p/biokanga/). Alignment allowed a mismatch rate of up to

10% of the length of the read. Only best unique hits were selected for differential expression analysis. The correlation between samples was performed using the Pearson's correlation on reads count. Unigenes were filtered to those with a sum of counts greater than or equal to 100 across samples. Read counts were normalized and tested for differential expression using the R DESeq package v1.12.0 with default parameters (Anders and Huber, 2010). 18 pairwise comparisons were completed (between genotypes at each time point and between time points for each genotype). Genes were considered DE for FDR ≤ 0.05 and $|\log_2(FC)| \geq \log_2(2)$. For genes that were found DE in one comparison, their differential expression was reconsidered in similar comparisons (e.g. genotype comparisons at 0 DPI), using FDR ≤ 0.075 and $|\log_2(FC)| \geq \log_2(1.8)$. Genes were considered expressed if more than 0.15 read per kilo base per million (RPKM) was detected. The chromosomal location of each Unigene was determined using BLASTN (best hit, evalue ≤ 10^{-5}) against the wheat GSS (http://www.wheatgenome.org). Unigenes putatively involved in hormonal pathways (PR genes) were recovered using tBLASTN from wheat protein sequences. RT-qPCR checks were performed using the Applied Biosystems 7900HT Fast Real-Time PCR System and results were analysed using R. The relative expression level was calculated by comparing the target gene with the reference Unigene Ta.2291 (ADP-ribosylation factor) using the $2^{-\Delta Cq}$ formula.

REFERENCES

Acharya, B.R., Raina, S., Maqbool, S.B., Jagadeeswaran, G., Mosher, S.L., Appel, H.M., Schultz, J.C., Klessig, D.F., and Raina, R. (2007). Overexpression of CRK13, an *Arabidopsis* cysteine-rich receptor-like kinase, results in enhanced resistance to *Pseudomonas syringae*: Overexpression of CRK13 induces pathogen defense. Plant J. 50: 488–499.

Allen, A.K., Neuberger, A., and Sharon, N. (1973). The purification, composition and specificity of wheat-germ agglutinin. Biochem J 131: 155–162.

An, C. and Mou, Z. (2011). Salicylic acid and its function in plant immunity: defense signal molecule salicylic acid. J. Integr. Plant Biol. 53: 412–428.

Anders, S. and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol 11: R106.

Ayliffe, M., Devilla, R., Mago, R., White, R., Talbot, M., Pryor, A., and Leung, H. (2011). Nonhost resistance of rice to rust pathogens. Mol. Plant. Microbe Interact. 24: 1143–1155.

Brueggeman, R., Rostoks, N., Kudrna, D., Kilian, A., Han, F., Chen, J., Druka, A., Steffenson, B., and Kleinhofs, A. (2002). The barley stem rust-resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases. Proc. Natl. Acad. Sci. 99: 9328–9333.

Catanzariti, A.-M., Dodds, P.N., and Ellis, J.G. (2007). Avirulence proteins from haustoria-forming pathogens: Avirulence proteins from haustoria-forming pathogens. FEMS Microbiol. Lett. 269: 181–188.

Chen, K., Du, L., and Chen, Z. (2003). Sensitization of defense responses and activation of programmed cell death by a pathogen-induced receptor-like protein kinase in Arabidopsis. Plant Mol. Biol. 53: 61–74.

Chen, K., Fan, B., Du, L., and Chen, Z. (2004). Activation of hypersensitive cell death by pathogen-induced receptor-like protein kinases from *Arabidopsis*. Plant Mol. Biol. 56: 271–283.

Conte, S.S. and Lloyd, A.M. (2011). Exploring multiple drug and herbicide resistance in plants—Spotlight on transporter proteins. Plant Sci. 180: 196–203.

Dardick, C., Chen, J., Richter, T., Ouyang, S., and Ronald, P. (2006). The rice kinase database. A phylogenomic database for the rice kinome. PLANT Physiol. 143: 579–586.

Dardick, C. and Ronald, P. (2006). Plant and animal pathogen recognition receptors signal through non-RD kinases. PLoS Pathog. 2: e2.

Desmond, O.J., Manners, J.M., Stephens, A.E., MacLean, D.J., Schenk, P.M., Gardiner, D.M., Munn, A.L., and Kazan, K. (2008). The Fusarium mycotoxin deoxynivalenol elicits hydrogen peroxide production, programmed cell death and defence responses in wheat. Mol. Plant Pathol. 9: 435–445.

Dodds, P.N. and Rathjen, J.P. (2010). Plant immunity: towards an integrated view of plant–pathogen interactions. Nat. Rev. Genet. 11: 539–548.

Dornez, E., Croes, E., Gebruers, K., De Coninck, B., Cammue, B., Delcour, J., and Courtin, C. (2010). Accumulated evidence substantiates a role for three classes of wheat xylanase inhibitors in plant defense. Crit. Rev. Plant Sci. 29: 244–264.

Du, L. and Chen, Z. (2000). Identification of genes encoding receptor-like protein kinases as possible targets of pathogen-and salicylic acid-induced WRKY DNA-binding proteins in *Arabidopsis*. Plant J. 24: 837–847.

Ederli, L., Madeo, L., Calderini, O., Gehring, C., Moretti, C., Buonaurio, R., Paolocci, F., and Pasqualini, S. (2011). The *Arabidopsis thaliana* cysteine-rich receptor-like kinase CRK20 modulates host responses to *Pseudomonas syringae* pv. tomato DC3000 infection. J. Plant Physiol. 168: 1784–1794.

Fu, Z.Q., Guo, M., Jeong, B., Tian, F., Elthon, T.E., Cerny, R.L., Staiger, D., and Alfano, J.R. (2007). A type III effector ADP-ribosylates RNA-binding proteins and quells plant immunity. Nature 447: 284–288.

Høj, P.B. and Fincher, G.B. (1995). Molecular evolution of plant β-glucan endohydrolases. Plant J. 7: 367–379.

Hurni, S., Brunner, S., Buchmann, G., Herren, G., Jordan, T., Krukowski, P., Wicker, T., Yahiaoui, N., Mago, R., and Keller, B. (2013). Rye *Pm8* and wheat *Pm3* are orthologous genes and show evolutionary conservation of resistance function against powdery mildew. Plant J. 76: 957–969.

Jain, M. (2011). Next-generation sequencing technologies for gene expression profiling in plants. Brief. Funct. Genomics.

Jung, K.-H., Cao, P., Seo, Y.-S., Dardick, C., and Ronald, P.C. (2010). The Rice Kinase Phylogenomics Database: a guide for systematic analysis of the rice kinase super-family. Trends Plant Sci. 15: 595–599.

Kang (2011). Plant ABC Transporters. Arab. Book.

Koga, H., Zeyen, R.J., Bushnell, W.R., and Ahlstrand, G.G. (1988). Hypersensitive cell death, autofluorescence, and insoluble silicon accumulation in barley leaf epidermal cells under attack by *Erysiphe graminis* f. sp. *hordei*. Physiol. Mol. Plant Pathol. 32: 395–409.

Kwak, K.J., Kim, J.Y., Kim, Y.O., and Kang, H. (2006). Characterization of Transgenic *Arabidopsis* Plants Overexpressing High Mobility Group B Proteins under High Salinity, Drought or Cold Stress. Plant Cell Physiol. 48: 221–231.

Lagaert, S., Beliën, T., and Volckaert, G. (2009). Plant cell walls: Protecting the barrier from degradation by microbial enzymes. Semin. Cell Dev. Biol. 20: 1064–1073.

Langlois-Meurinne, M. (2005). Pathogen-responsive expression of glycosyltransferase genes *UGT73B3* and *UGT73B5* is necessary for resistance to *Pseudomonas syringae* pv tomato in *Arabidopsis*. PLANT Physiol. 139: 1890–1901.

Lee, H.J., Kim, J.S., Yoo, S.J., Kang, E.Y., Han, S.H., Yang, K.-Y., Kim, Y.C., McSpadden Gardener, B., and Kang, H. (2012). Different roles of glycine-rich RNA-binding protein7 in plant defense against *Pectobacterium carotovorum*, *Botrytis cinerea*, and tobacco mosaic viruses. Plant Physiol. Biochem. 60: 46–52.

Liu, B., Xue, X., Cui, S., Zhang, X., Han, Q., Zhu, L., Liang, X., Wang, X., Huang, L., Chen, X., and Kang, Z. (2010). Cloning and characterization of a wheat β-1,3-glucanase gene induced by the stripe rust pathogen *Puccinia striiformis* f. sp. *tritici*. Mol. Biol. Rep. 37: 1045–1052.

Lorang, J.M., Sweat, T.A., and Wolpert, T.J. (2007). Plant disease susceptibility conferred by a "resistance" gene. Proc. Natl. Acad. Sci. 104: 14861–14866.

Marone, D., Russo, M., Laidò, G., De Leonardis, A., and Mastrangelo, A. (2013). Plant nucleotide binding site–leucine-rich repeat (NBS-LRR) genes: active guardians in host defense responses. Int. J. Mol. Sci. 14: 7302–7326.

Matsushita, A., Inoue, H., Goto, S., Nakayama, A., Sugano, S., Hayashi, N., and Takatsuji, H. (2013). Nuclear ubiquitin proteasome degradation affects WRKY45 function in the rice defense program. Plant J. 73: 302–313.

McIntosh, R.A., Zhang, P., Cowger, C., Parks, R., Lagudah, E.S., and Hoxha, S. (2011). Rye-derived powdery mildew resistance gene *Pm8* in wheat is suppressed by the *Pm3* locus. Theor. Appl. Genet. 123: 359–367.

Mendgen, K. (1997). The Uredinales. In The Mycota V Part B (Springer-Verlag), pp. 79–94.

Meyberg, M. (1988). Selective staining of fungal hyphae in parasitic and symbiotic plant-fungus associations. Histochemistry 88: 197–199.

Molina, A., Görlach, J., Volrath, S., and Ryals, J. (1999). Wheat genes encoding two types of PR-1 proteins are pathogen inducible, but do not respond to activators of systemic acquired resistance. Mol. Plant. Microbe Interact. 12: 53–58.

Nelson, D. and Werck-Reichhart, D. (2011). A P450-centric view of plant evolution: P450-centric evolution. Plant J. 66: 194–211.

Nirmala, J., Drader, T., Chen, X., Steffenson, B., and Kleinhofs, A. (2010). Stem rust spores elicit rapid RPG1 phosphorylation. Mol. Plant. Microbe Interact. 23: 1635–1642.

Nirmala, J., Drader, T., Lawrence, P.K., Yin, C., Hulbert, S., Steber, C.M., Steffenson, B.J., Szabo, L.J., von Wettstein, D., and Kleinhofs, A. (2011). Concerted action of two avirulent spore effectors activates Reaction to Puccinia graminis 1 (Rpg1)-mediated cereal stem rust resistance. Proc. Natl. Acad. Sci. 108: 14676–14681.

Paolacci, A.R., Tanzarella, O.A., Porceddu, E., and Ciaffi, M. (2009). Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. BMC Mol. Biol. 10: 11.

Park, H.-J., Kwon, C.-S., Woo, J.-Y., Lee, G.-J., Kim, Y.-J., and Paek, K.-H. (2011). Suppression of UDP-glycosyltransferase-coding *Arabidopsis thaliana UGT74E2* gene expression leads to increased resistance to *Pseudomonas syringae* pv. tomato DC3000 infection. Plant Pathol. J. 27: 170–182.

Rafiqi, M., Ellis, J.G., Ludowici, V.A., Hardham, A.R., and Dodds, P.N. (2012). Challenges and progress towards understanding the role of effectors in plant–fungal interactions. Curr. Opin. Plant Biol. 15: 477–482.

Rayapuram, C., Jensen, M.K., Maiser, F., Shanir, J.V., HornshøJ, H., Rung, J.H., Gregersen, P.L., Schweizer, P., Collinge, D.B., and Lyngkjaer, M.F. (2012). Regulation of basal resistance by a powdery mildew-induced cysteine-rich receptor-like protein kinase in barley: CRK-regulated basal resistance in barley. Mol. Plant Pathol. 13: 135–147.

Rostoks, N., Steffenson, B.J., and Kleinhofs, A. (2004). Structure and expression of the barley stem rust resistance gene *Rpg1* messenger RNA. Physiol. Mol. Plant Pathol. 64: 91–101.

Schie, C.C.N., Haring, M.A., and Schuurink, R.C. (2007). Tomato linalool synthase is induced in trichomes by jasmonic acid. Plant Mol. Biol. 64: 251–263.

Schuhegger, R., Nafisi, M., Mansourova, M., Petersen, B.L., Olsen, C.E., Svatos, A., Halkier, B.A., and Glawischnig, E. (2006). CYP71B15 (PAD3) catalyzes the final step in camalexin biosynthesis. PLANT Physiol. 141: 1248–1254.

Shiu, S.-H. (2004). Comparative analysis of the receptor-like kinase family in arabidopsis and rice. PLANT CELL ONLINE 16: 1220–1234.

Singh, D., Bhaganagare, G., Bandopadhyay, R., Prabhu, K.V., Gupta, P.K., and Mukhopadhyay, K. (2012). Targeted spatio-temporal expression based characterization of state of infection and time-point of maximum defense in wheat NILs during leaf rust infection. Mol. Biol. Rep. 39: 9373–9382.

Strober, W. (2001). Trypan blue exclusion test of cell viability. In current protocols in immunology (John Wiley & Sons, Inc.), p. A.3B.1–A.3B.2.

Sweat, T.A., Lorang, J.M., Bakker, E.G., and Wolpert, T.J. (2008). Characterization of natural and induced variation in the *LOV1* gene, a CC-NB-LRR gene conferring victorin sensitivity and disease susceptibility in Arabidopsis. Mol. Plant. Microbe Interact. 21: 7–19.

Taniguchi, S., Hosokawa-Shinonaga, Y., Tamaoki, D., Yamada, S., Akimitsu, K., and Gomi, K. (2013). Jasmonate induction of the monoterpene linalool confers resistance to rice bacterial blight and its biosynthesis is

regulated by JAZ protein in rice: Role of volatile on jasmonate-induced resistance. Plant Cell Environ. 37: 451-61.

Voegele, R.T., Hahn, M., and Mendgen, K. (2009). The Uredinales: cytology, biochemistry, and molecular biology. In The Mycota, H.B. Deising, ed (Springer Berlin Heidelberg: Berlin, Heidelberg), pp. 69–98.

Voegele, R.T. and Mendgen, K.W. (2011). Nutrient uptake in rust fungi: how sweet is parasitic life? Euphytica 179: 41–55.

Wasternack, C. and Hause, B. (2013). Jasmonates: biosynthesis, perception, signal transduction and action in plant stress response, growth and development. An update to the 2007 review in Annals of Botany. Ann. Bot. 111: 1021–1058.

Wrzaczek, M., Brosché, M., Salojärvi, J., Kangasjärvi, S., Idänheimo, N., Mersmann, S., Robatzek, S., Karpiński, S., Karpińska, B., and Kangasjärvi, J. (2010). Transcriptional regulation of the CRK/DUF26 group of receptor-like protein kinases by ozone and plant hormones in *Arabidopsis*. BMC Plant Biol. 10: 95.

Xiang, Y., Song, M., Wei, Z., Tong, J., Zhang, L., Xiao, L., Ma, Z., and Wang, Y. (2011). A jacalin-related lectin-like gene in wheat is a component of the plant defence system. J. Exp. Bot. 62: 5471–5483.

Xin, Z., Wang, Q., Yu, Z., Hu, L., Li, J., Xiang, C., Wang, B., and Lou, Y. (2014). Overexpression of a Xylanase Inhibitor Gene, OsHI-XIP, Enhances Resistance in Rice to Herbivores. Plant Mol. Biol. Report. 32: 465–475.

Ziemienowicz, A., Haasen, D., Staiger, D., and Merkle, T. (2003). Arabidopsis transportin1 is the nuclear import receptor for the circadian clock-regulated RNA-binding protein AtGRP7. Plant Mol. Biol. 53: 201–212.

SUPPLEMENTAL

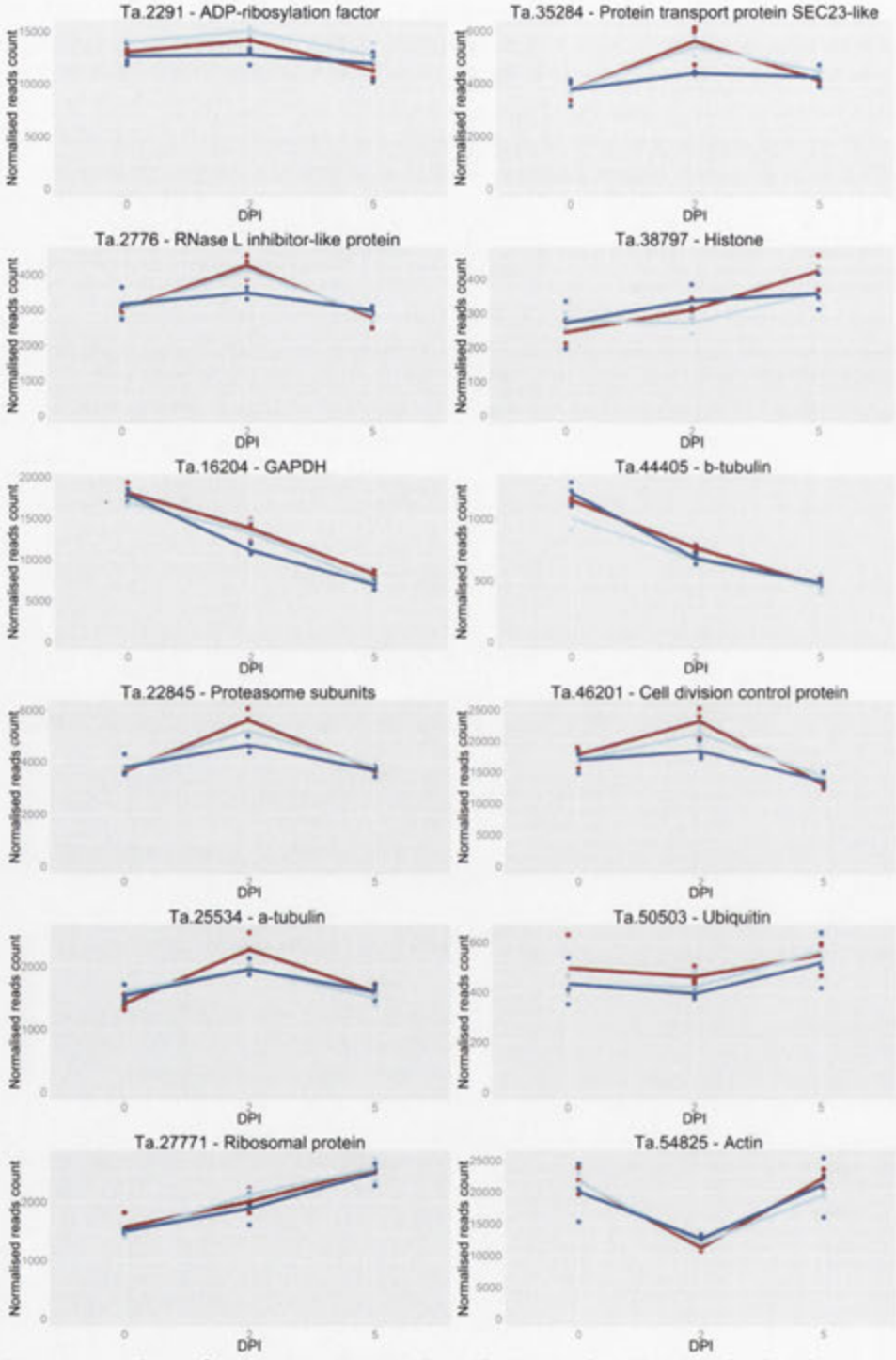
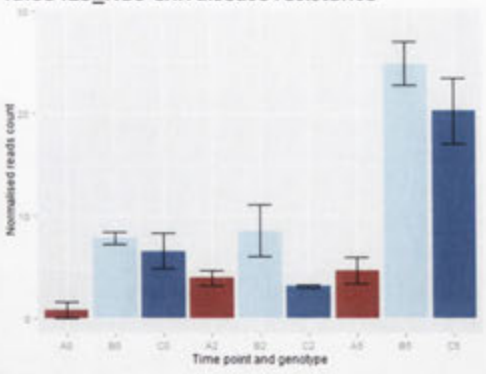
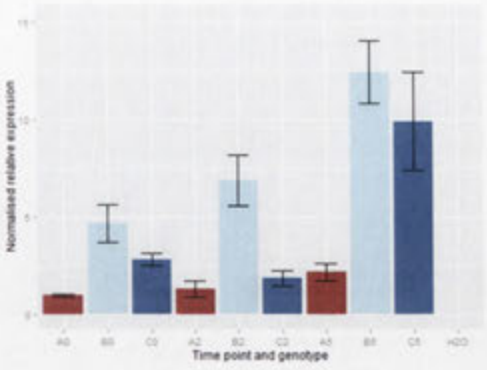


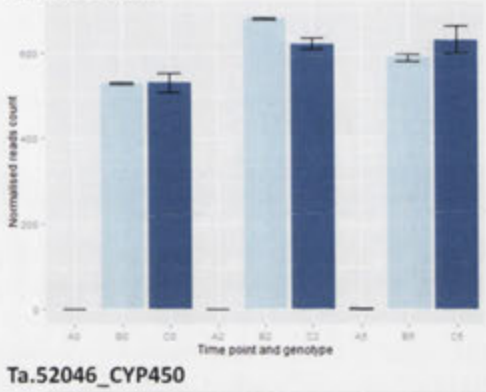
Figure Sup. 1: Wheat reference gene expression. For each Unigene, the DESeq-normalised read count is reported for each sample (dots) at the different time points. Lines represent the biological replicates average. Colours represent genotypes: Col (red), Col-NS765 (pale blue), Col-NS766 (dark blue).

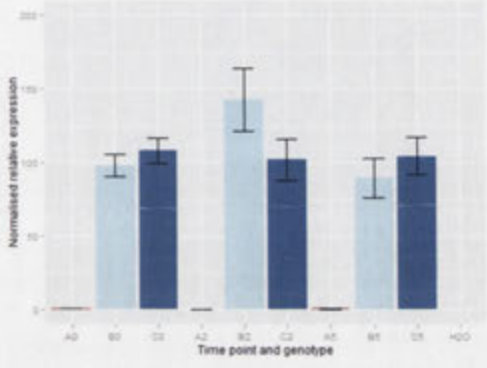
Ta.88423_NBS-LRR disease resistance



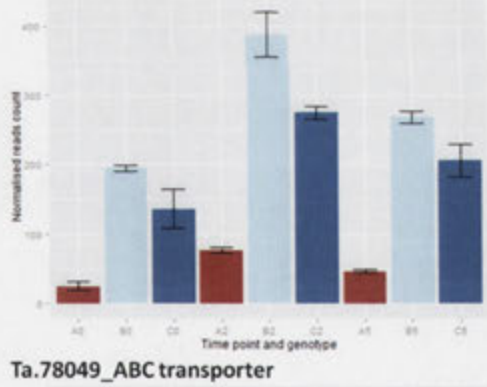


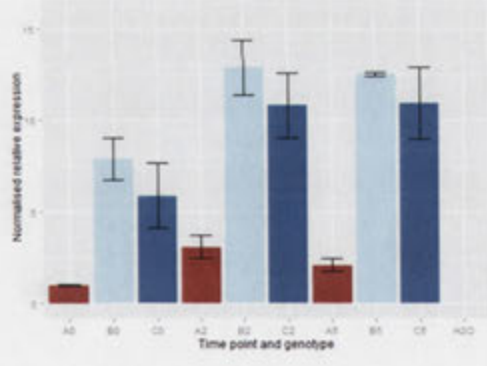
Ta.104804_CRK



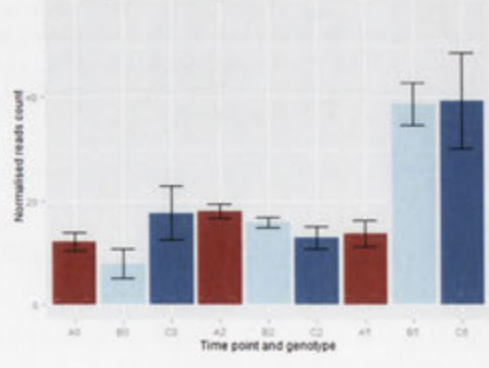


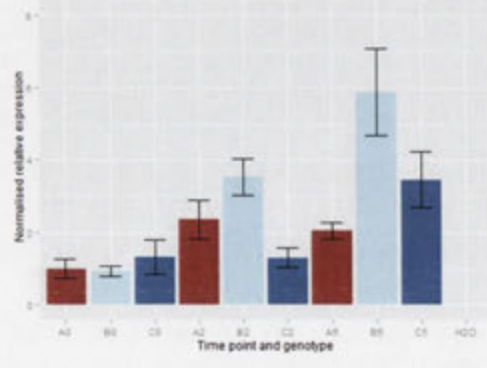
Ta.52046_CYP450





Ta.78049_ABC transporter





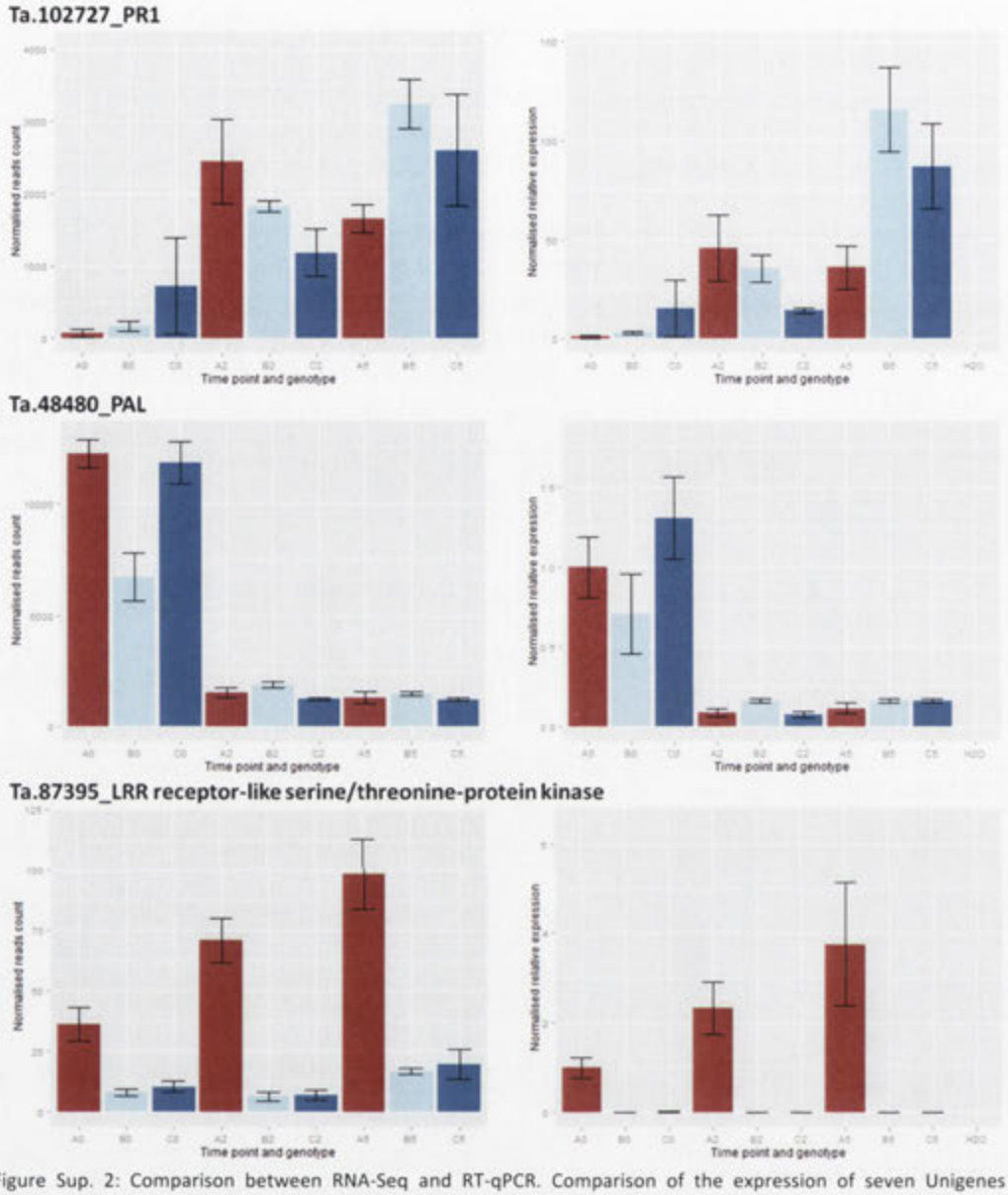


Figure Sup. 2: Comparison between RNA-Seq and RT-qPCR. Comparison of the expression of seven Unigenes determined through RNA-Seq or RT-qPCR. Left panels) Means of the DESeq-normalised reads count from RNA-Seq for each genotype at each time point. Right panels) Means of the relative expression from RT-qPCR, normalised to Col at 0 DPI (A0). Letters and colours represent genotypes (A: Col, red; B: Col-NS765, pale blue; C: Col-NS766, dark blue) and numbers time points (0: 0 DPI, 2: 2 DPI, 5: 5 DPI). Errors bars represent the standard error of the mean. For each gene, results from RNA-Seq and RT-qPCR are very similar and follow the same trend.

File Sup. 1: Genes of interest. Details about the Unigenes found differentially expressed, through RNA-Seq, between Col and both Col-NS765 and Col-NS766, at 0, 2 and 5 DPI.

File Sup. 2: RNA-Seq alignment count. Count of reads that aligned to each Unigene for all 27 samples used in RNA-Seq.

File Sup. 3: RNA-Seq normalised count. Count of reads that aligned to the 67,156 Unigene used in the differential expression analysis, after DESeq normalisation.

Chapter 5

General Discussion

List of major abbreviations

7DL-Sup Canthatch 7DL suppressor; **9k chip** 9k SNP chip; **BSA** bulked segregant analysis; **Col** Columbus; **Col-NS** Columbus-NonSuppressor; **CRKs** cysteine-rich receptor-like kinases; **CTH** Canthatch; **CTH-NS** Canthatch-NonSuppressor; **DE** differentially expressed; **DPI** days post-inoculation; **ETI** effector-triggered immunity; **GBS** genotyping by sequencing; **NS** non-suppressor; **Pgt** *Puccinia graminis* f. sp. *tritici*; **PTI** pattern-triggered immunity; **R-genes** race-specific genes; **RAD-Seq** restriction-site associated DNA sequencing; **Tc** Thatcher.

Background

Wheat is one of the most important food crops in the world, and contributes to a substantial part of human nutrition. Maintaining good production is crucial for preserving global food security. Like any other plants, wheat health can be greatly affected by pathogens, notably rust fungi, which cause some of the most destructive plant diseases and are a constant threat to wheat production. Particularly, wheat stem rust (*Puccinia graminis* f. sp. *tritici*, abbreviated as *Pgt*) has always been a major concern due to its rapid spread on a very large scale and its capability of ruining an entire healthy crop within a month. A highly virulent pathotype of *Pgt*, Ug99 (also known as TTKSK), was first observed in Uganda in 1999 (Pretorius et al., 2000), and has now spread to many regions from South Africa to Iran (Nazari et al., 2009; Pretorius et al., 2010; Wanyera et al., 2006). It is predicted that Ug99 and its variants will migrate to North Africa, the Middle East, Asia and beyond, menacing global food security. Controlling this disease is therefore critical for sustaining food production.

The most successful and economic way to control stem rust is the use of genetic resistance in wheat cultivars. All named stem rust resistance genes (R-genes), except *Sr2*, are race-specific and only provide resistance against a small range of stem rust races. Several R-genes are effective against Ug99 (Singh et al., 2011), and international collaborations are now being made to appropriately release resistant cultivars in the affected regions, which are expected to greatly reduce the threat posed by Ug99. However, only two of these genes, *Sr33* and *Sr35*, have been cloned so far (Periyannan et al., 2013; Saintenac et al., 2013). Cloning R-genes permits to efficiently introgress them into cultivars of agronomic importance. In addition, their study can reveal the mechanism behind the resistance, which can help in developing new strategies for fighting rust. Nevertheless, rust pathogens are known for their ability to rapidly overcome R-genes, and constant efforts to find new sources of resistance

must remain a priority. One way to do that consists in characterizing suppressors of resistance. Indeed, there are many cases of resistance genes that appeared to be non-functional when transferred to a different genetic background, especially in hexaploid wheat when genes are brought from a lower ploidy species. This phenomenon seems often to be due to the presence of suppressors that inhibit the expression of resistance.

This study first aimed to genetically fine map the Canthatch (CTH) suppressor located on the chromosome 7DL (7DL-Sup) (Kerber and Green, 1980), as the genes it inhibits were found to confer robust resistance against Ug99 (Lagudah and Jin, unpublished). The 7DL-Sup seems quite common in wheat (Kerber, 1983), and being able to inactivate it would facilitate the use of the corresponding R-genes. The biological resources already available were expected to be of great help in order to clone the 7DL-Sup. These resources included several mutants of the 7DL-Sup that are resistant to stem rust (Kerber, 1991; Williams et al., 1992), as well as backcrossed lines for which the non-suppressor (NS) allele was inferred to be introgressed into Columbus (Col) (Kerber, 1991). These lines, namely Col-NS765 and Col-NS766 (both are referred as Col-NS), were of particular interest as Col, being closely related to CTH and susceptible to stem rust, was thought to possess a common genetic background, including the 7DL-Sup, while still displaying polymorphism adequate for mapping. A $F_{2:3}$ mapping population descended from the cross between Col and Col-NS766 was thus developed, in order to map the 7DL-Sup, but was actually segregating for a locus of resistance on chromosome 7AL. This population was phenotyped using two strains of stem rust, Ug99 and #313, and several sets of molecular markers were mapped. These markers were identified though various approaches; markers already known in the literature, the wheat 9k SNP chip (9k chip) and genotyping by sequencing (GBS) using the restriction-site associated DNA sequencing (RAD-Seq) method in association with bulked segregant analysis (BSA). In addition to the mapping, the resistance in Col-NS was characterized through histological observations and transcriptomic analyses using RNA sequencing (RNA-Seq).

The resistance response

Phenotyping of the Col x Col-NS766 $F_{2:3}$ population for resistance against Ug99 and #313 resulted in similar observations, indicating that the resistance involved against both races is conferred by the same locus. This resistance was first observable in Col-NS at 2 days post-inoculation (DPI), at least against #313, and consisted in the death of cells that are adjacent to an infection site, that is, where the appressorium was formed. This usually completely stopped

the pathogen infection. Cell death is usually the sign of effector-triggered immunity (ETI), but an enhanced response of pattern-triggered immunity (PTI) cannot be excluded.

At the transcriptome level, RNA-Seq analysis resulted in the discovery of 353 genes that were differentially expressed (DE) between Col and Col-NS. Although many of these genes are expected to be DE due to allelic differences between Col and Col-NS (i.e. due to the introgression of CTH genomic regions in Col-NS), they could all be involved in the resistance. Some of these genes even presented annotations strongly suggesting a role in plant defence. Functional analysis of these genes should now be undertaken in order to confirm their role in the resistance.

A resistance locus on chromosome 7AL

In the course of genetic mapping of the locus responsible for the resistance in Col-NS766, it became clear that this locus was not on chromosome 7DL, as it was expected, but on 7AL. This was first realized with the detection, through BSA using RAD-Seq, of hundreds of markers putatively linked to the stem rust resistance that were found *in silico* located on chromosome 7AL. All linked markers were later confirmed to be located on 7AL using nullisomic lines. Moreover, for all the markers linked to stem rust resistance, the independent backcrossed line Col-NS765 was always showing the same allele as Col-NS766, indicating that this locus is essential for resistance. This was also shown through whole transcriptome analysis using RNA-Seq, where except for chromosome 6A, most genes consistently DE between Col and Col-NS were located on 7AL.

There is clear evidence in the literature for the presence of a suppressor of stem rust resistance on chromosome 7DL in CTH (Kerber, 1983, 1991; Kerber and Aung, 1995; Kerber and Green, 1980; Williams et al., 1992). Therefore, the simplest explanation for linkage of resistance with an introgressed region on chromosome 7AL in Col-NS, is that Col possesses neither the 7DL-Sup, as it was previously thought, nor the suppressed resistance genes. Thus, only the introgression of R-genes, presumably on 7AL, could confer resistance to stem rust in Col-NS. This was considered unlikely since, like CTH, Col is closely related to Thatcher (Tc), which is most likely the donor of the 7DL-Sup (Kerber, 1991). Col was developed by backcrossing six times RL-4137, which shares approximately half of the Tc genome, into the recurrent parent Neepawa, which may share more than 90 % of the Tc genome (Campbell and Czarnecki, 1981) (Figure 1). However, it is possible that the 7DL-Sup and the 7AL resistance locus were

not retained during Col selection. Indeed, this line was selected for the pre-harvest sprouting resistance trait from RL-4137. The genes conferring this trait may be located on 7DL and/or 7AL, which would lead to the selection of these chromosomes from RL-4137 instead of Tc. Interestingly, QTLs of resistance to pre-harvest sprouting have been found on chromosome 7A and 7D in the wheat line W98646, for which one of the parents is RL-4137 (Singh, 2008). The 7DL-Sup and/or the 7AL locus could also have been lost in Neepawa. Although very unlikely, it is also possible that the donor of the 7DL-Sup in CTH is not Tc, but the other parent, Kenya-farmer. On the other hand, the 7AL locus most likely originated from Tc, as all the markers on 7AL that were tested on Tc and found linked to stem rust resistance, were showing the same allele in Col-NS, CTH and Tc.

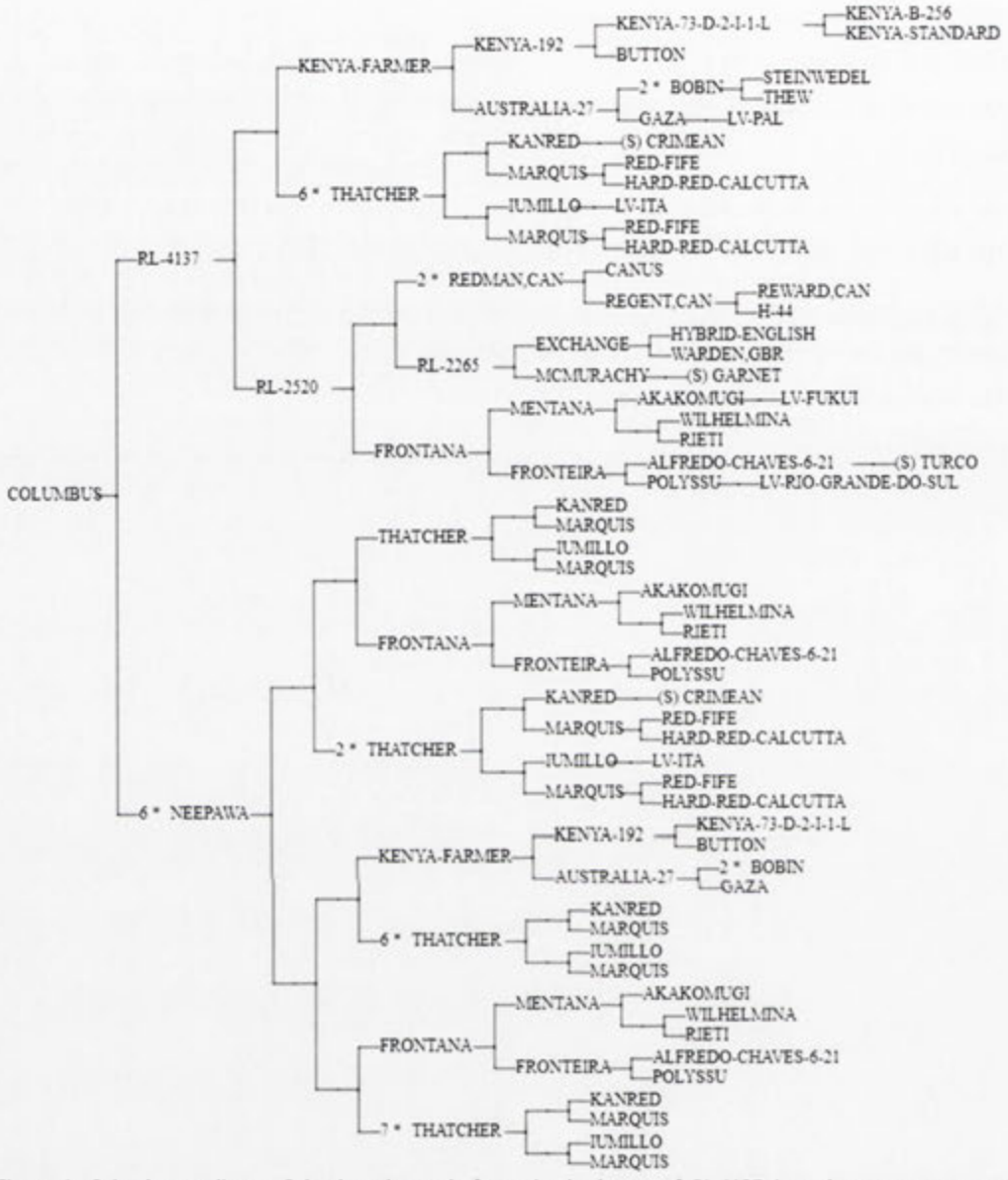


Figure 1: Columbus pedigree. Columbus descends from the backcross of RL-4137 into the recurrent parent Neepawa, which are both related to Thatcher. Canthatch has a similar pedigree to one of Neepawa's relatives (THATCHER*6/KENYA-FARMER). Source: Genetic Resources Information System for Wheat and Triticale (http://wheatpedigree.net/).

Kerber originally demonstrated the presence of the 7DL-Sup by crossing several common wheat cultivars with CTH nullisomic 7D (Kerber, 1983). To test for the presence of the 7DL-Sup in Col, a similar approach was taken. Crosses were performed between Col and CTH nullisomic 7D, as well as with the CTH ditelosomic 7DS, but was unsuccessful as the discrimination between resistant and susceptible responses, with stem rust race #313, was too difficult. Crossing Col with CTH could also reveal the presence of the 7DL-Sup in Col, if resistance was

seen in the progeny, which would be unlikely if both Col and CTH possessed the 7DL-Sup. This cross was also done using CTH-K, and although F_2 lines were observed with higher resistance than the parents, once again the phenotype was not clear enough to confirm that the 7DL-Sup was absent in Col.

On the other hand, Col is not completely susceptible to Ug99 and #313, and for the latter, histological observations showed that the partial resistance in Col may share some features with the more robust resistance of Col-NS. Therefore, some of the genes involved in the resistance in Tc, may have been passed on to Col, and possibly more from other genotypes, since the resistance in genotypes with Col background is better than that in genotypes with CTH background. Furthermore, Col may possess the R-gene responsible for the pathogen recognition, but may be missing downstream components, essential for carrying out an effective defence response.

RNA-Seq data shed light on several genes coding for cysteine-rich receptor-like kinases (CRKs), which were consistently DE between Col and Col-NS and which mapped to 7AL. Although CRKs, like most receptor-like kinases (RLKs) are known for having an extracellular domain and acting as receptors (Chen et al., 2003), the putative orthologues in rice of the wheat CRK genes on 7AL all belong to the same sub-family of receptor-like cytoplasmic kinases (RLCK-OS2), which are predicted to be cytoplasmic and could interact with RLKs to relay intracellular signalling (Lin et al., 2013). They could also act as guardees or interact with them in ETI. In any case, it would be of great interest to undertake functional analyses of these genes.

Multigenic resistance

Although the segregation ratio, resistant:segregating:susceptible, for the $F_{2:3}$ families was not statistically different from a one-gene segregation (1:2:1), the reaction type of individual F_3 plants that came from families considered non-segregating, was often not homogeneous. This variability could possibly be attributed to the environment, which was observed to sometimes affect the stem rust infection. However, variation could also be due to the segregation of multiple genes. In addition, segregation ratios in F_3 plants that came from families considered to be segregating for resistance, rarely fitted a one-gene segregation ratio (1:2:1, 3:1 or 1:3). Although the F_3 population size in a particular experiment may have been

too small for such statistics (15 to 20 plants), the variation in reaction type in these segregating families supports the hypothesis than more than one gene is segregating.

Moreover, mapping the 7AL resistance locus did not permit to completely associate a region with the resistance, although the genetic map obtained indicated that the cluster of co-segregating markers was the most likely position for such region. Here also, these data seem to indicate that multiple segregating genes could be involved in the resistance. The locus on 7AL would be essential, as it explains most of the phenotypes, and because resistance was never seen without this locus originating from the resistant parent Col-NS766. Finally, the retention of a large fragment, originating from CTH, on chromosome 6AL in the two independent backcrossed lines, Col-NS765 and Col-NS766, clearly indicates that this region was actively selected. Because these lines were selected for resistance to stem rust, one or several genes present on this chromosome have most likely a role in this resistance. These genes would, therefore, segregate in the mapping population, explaining the phenotypic variations and the atypical map obtained on 7AL. The gene(s) responsible for stem rust resistance were inherited from CTH, which is closely related to Tc. Tc is known for carrying many resistance genes, categorized or not, including some that are effective against Ug99 (Jin, 2007). However, the inheritance of these genes from Tc is also known to be complex, and several studies mentioned the possible involvement of multiple complementary genes and suppressors (Jin, 2007; Knott, 2000, 2001).

As depicted in chapter 3, several models of multigenic resistance could fit the phenotypic observations, but a bigger population is needed for determining their relevance. In addition, other populations, segregating for only one of these genes, are needed in order to individually map them and to understand their contribution to the resistance. This would notably be of great interest for the 7AL locus, in order to determine if it can alone provide some resistance, or if the other genes are needed. Indeed, instead of being involved in the resistance, these other genes could suppress it and were thus negatively selected in Col-NS. Nevertheless, this study indicates that the 7AL locus is a major determinant of resistance, for which several robust markers were developed. These markers will be useful tools in deployment of this resistance in wheat breeding.

REFERENCES

Campbell, A.B., and Czarnecki, E.M. (1981). Columbus hard red spring wheat. Can. J. Plant Sci. *61*, 147–148.

Chen, K., Du, L., and Chen, Z. (2003). Sensitization of defense responses and activation of programmed cell death by a pathogen-induced receptor-like protein kinase in *Arabidopsis*. Plant Mol. Biol. *53*, 61–74.

Jin, Y. (2007). Resistance to race TTKS of *Puccinia graminis* f. sp. *tritici* in Chris and related spring wheat. Phytopathology *97*, S162–S162.

Kerber, E.R. (1983). Suppression of rust resistance in amphiploids of *Triticum*. In Proc. 6th International Wheat Genetics Symposium, (Kyoto, Japan), pp. 813–817.

Kerber, E.R. (1991). Stem-rust resistance in "Canthatch" hexaploid wheat induced by a nonsuppressor mutation on chromosome 7DL. Genome *34*, 935–939.

Kerber, E.R., and Aung, T. (1995). Confirmation of nonsuppressor mutation of stem rust resistance in "Canthatch"common wheat. Crop Sci. *35*, 743–744.

Kerber, E.R., and Green, G.J. (1980). Suppression of stem rust resistance in the hexaploid wheat cv. Canthatch by chromosome 7DL. Can. J. Bot. *58*, 1347–1350.

Knott, D.R. (2000). The inheritance of stem rust resistance in Thatcher wheat. Can. J. Plant Sci. *80*, 53–63.

Knott, D.R. (2001). The relationship between seedling and field resistance to two races of stem rust in Thatcher wheat. Can. J. Plant Sci. *81*, 415–418.

Lin, W., Ma, X., Shan, L., and He, P. (2013). Big roles of small kinases: The complex functions of receptor-like cytoplasmic kinases in plant immunity and development: Multiple functions of RLCKs. J. Integr. Plant Biol. *55*, 1188–1197.

Nazari, K., Mafi, M., Yahyaoui, A., Singh, R.P., and Park, R.F. (2009). Detection of wheat stem rust (*Puccinia graminis* f. sp. *tritici*) race TTKSK (Ug99) in Iran. Plant Dis. *93*, 317–317.

Periyannan, S., Moore, J., Ayliffe, M., Bansal, U., Wang, X., Huang, L., Deal, K., Luo, M., Kong, X., Bariana, H., et al. (2013). The gene *Sr33*, an ortholog of barley *Mla* genes, encodes resistance to wheat stem rust race Ug99. Science *341*, 786–788.

Pretorius, Z.A., Singh, R.P., Wagoire, W.W., and Payne, T.S. (2000). Detection of virulence to wheat stem rust resistance gene *Sr31* in *Puccinia graminis* . f. sp. *tritici* in Uganda. Plant Dis. *84*, 203–203.

Pretorius, Z.A., Bender, C.M., Visser, B., and Terefe, T. (2010). First report of a *Puccinia graminis* f. sp. *tritici* race virulent to the *Sr24* and *Sr31* wheat stem rust resistance genes in South Africa. Plant Dis. *94*, 784–784.

Saintenac, C., Zhang, W., Salcedo, A., Rouse, M.N., Trick, H.N., Akhunov, E., and Dubcovsky, J. (2013). Identification of wheat gene *Sr35* that confers resistance to Ug99 stem rust race group. Science *341*, 783–786.

Singh, R. (2008). Identification and validation of genomic regions associated with pre-harvest sprouting resistance in white-grained wheat (*Triticum aestivum* L.). University of Saskatchewan.

Singh, R.P., Hodson, D.P., Huerta-Espino, J., Jin, Y., Bhavani, S., Njau, P., Herrera-Foessel, S., Singh, P.K., Singh, S., and Govindan, V. (2011). The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. Annu. Rev. Phytopathol. *49*, 465–481.

Wanyera, R., Kinyua, M.G., Jin, Y., and Singh, R.P. (2006). The spread of stem rust caused by *Puccinia graminis* f. sp. *tritici* , with virulence on *Sr31* in wheat in eastern africa. Plant Dis. *90*, 113–113.

Williams, N.D., Miller, J.D., and Klindworth, D.L. (1992). Induced mutations of a genetic suppressor of resistance to wheat stem rust. Crop Sci. *32*, 612–616.